

2

UTILITY FILE COPY

AD-A203 706



TECHNICAL REPORT RD-RE-87-3

**DIAGNOSTIC PARAMETER DETERMINATION
FOR A CLASS OF THREE-PARAMETER
PROBABILITY DISTRIBUTIONS**

S. H. Lehnigk and H. P. Dudel
Research Directorate
Research, Development, & Engineering Center

JULY 1987



U.S. ARMY MISSILE COMMAND

Redstone Arsenal, Alabama 35898-5000

Approved for public release; distribution unlimited.

DTIC
SELECTE
7 Feb 1989
S
2
E
D

DISPOSITION INSTRUCTIONS

**DESTROY THIS REPORT WHEN IT IS NO LONGER NEEDED. DO NOT
RETURN IT TO THE ORIGINATOR.**

DISCLAIMER

**THE FINDINGS IN THIS REPORT ARE NOT TO BE CONSTRUED AS AN
OFFICIAL DEPARTMENT OF THE ARMY POSITION UNLESS SO DESIGNATED BY OTHER AUTHORIZED DOCUMENTS.**

TRADE NAMES

**USE OF TRADE NAMES OR MANUFACTURERS IN THIS REPORT DOES
NOT CONSTITUTE AN OFFICIAL INDORSEMENT OR APPROVAL OF
THE USE OF SUCH COMMERCIAL HARDWARE OR SOFTWARE.**

REPORT DOCUMENTATION PAGE				Form Approved OMB No 0704-0188 Exp Date Jun 30, 1986	
1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED			1b. RESTRICTIVE MARKINGS		
2a. SECURITY CLASSIFICATION AUTHORITY			3. DISTRIBUTION/AVAILABILITY OF REPORT		
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE			Approved for public release; distribution unlimited.		
4. PERFORMING ORGANIZATION REPORT NUMBER(S) RD- RE 87-3			5. MONITORING ORGANIZATION REPORT NUMBER(S)		
6a. NAME OF PERFORMING ORGANIZATION Research Directorate RD&E Center		6b. OFFICE SYMBOL (If applicable) AMSMI-RD-RE-OP	7a. NAME OF MONITORING ORGANIZATION		
6c. ADDRESS (City, State, and ZIP Code) Commander, US Army Missile Command ATTN: AMSMI-RD-RE-OP Redstone Arsenal, AL 35898-5248			7b. ADDRESS (City, State, and ZIP Code)		
8a. NAME OF FUNDING/SPONSORING ORGANIZATION		8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER		
8c. ADDRESS (City, State, and ZIP Code)			10. SOURCE OF FUNDING NUMBERS		
			PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.
					WORK UNIT ACCESSION NO
11. TITLE (Include Security Classification) Diagnostic Parameter Determination for a Class of Three-Parameter Probability Distributions (U)					
12. PERSONAL AUTHOR(S) S. H. Lehnigk and H. P. Dudel					
13a. TYPE OF REPORT Final Technical		13b. TIME COVERED FROM _____ TO Sep 86		14. DATE OF REPORT (Year, Month, Day) July 1987	
				15. PAGE COUNT 45	
16. SUPPLEMENTARY NOTATION					
17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)		
FIELD	GROUP	SUB-GROUP	Hyper-Gamma Distribution, Three Parameters, Density Estimation, Approximation of Cumulative Distribution		
19. ABSTRACT (Continue on reverse if necessary and identify by block number)					
<p>The problem of parameter estimation of the non-shifted hyper-Gamma class is being considered. The approach is based on an applicability criterion which provides the opportunity to determine the parameters by means of three equations derived from the first and second moments and an analytical approximation of the logarithm of the cumulative distribution function. The parameter estimation process requires the iterative solution of two equations. Examples are given to verify the efficiency of the proposed parameter determination method. Recently obtained results on maximum-likelihood parameter estimation for the hyper-Gamma class may turn out to be practically more reliable than those presented in this report, especially for non-smooth data.</p>					
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT <input checked="" type="checkbox"/> UNCLASSIFIED/UNLIMITED <input type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS			21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED		
22a. NAME OF RESPONSIBLE INDIVIDUAL S. H. Lehnigk			22b. TELEPHONE (Include Area Code) (205) 876-3526		22c. OFFICE SYMBOL AMSMI-RD-RD-OP

TABLE OF CONTENTS

	<u>Page</u>
I. INTRODUCTION	1
II. NOTATIONS AND FORMULAS	4
III. THE APPLICABILITY CRITERION	6
IV. DETERMINATION OF THE PARAMETERS	8
V. THE ITERATION PROCESS	13
VI. EXAMPLES	15
REFERENCES	26
APPENDIX A. THE EQUATION $h(\rho, \sigma) = 0$	A-1
APPENDIX B. APPROXIMATION OF $v(u)$ by $v^*(u)$	B-1
APPENDIX C. THE COEFFICIENT σ AS A FUNCTION OF β	C-1
APPENDIX D. CONVERGENCE OF THE ITERATION PROCESS	D-1

LIST OF TABLES

<u>Figure</u>	<u>Title</u>	<u>Page</u>
1	Empirical Example #3 - Absolute Frequencies.....	9
2	Empirical Example #3 - Relative Frequencies.....	10
3	SYNEX #8: Exponential Weibull Distribution (B=1, P=0, Beta=1).....	16
4	SYNEX #9: Weibull Distribution (B=1, P=-1, Beta=2).....	17
5	SYNEX #10: Weibull Distribution (B=2, P=-2, Beta=3).....	18
6	SYNEX #10: Weibull Distribution (B=2, P=12, Beta=3).....	19
7	SYNEX #11: Weibull Distribution (B=5, P=0.5, Beta=0.5).....	20
8	Empirical Example #3.....	21

Accession For	
NTIS (GSA&I)	<input checked="" type="checkbox"/>
DTIC T/R	<input checked="" type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Availability for	
Dist	Special
A-1	



SUMMARY

The application of a class of continuous, one-sided, three-parameter probability distributions is being considered. The parameters represent scale and initial and terminal shape of the associated probability density function. The class contains as special cases (for specific numerical values of the shape parameters) the following well-known distributions: Gauss, Weibull, exponential, Rayleigh, Gamma, chi-square, Maxwell, and Wien. The objective is to present and discuss a parameter determination technique which uses cumulative frequency data. The approach is based on an applicability criterion for the considered distribution class which provides the opportunity to determine the parameter values by means of three equations derived from the first and second moments, and an analytical approximation of the logarithm of the cumulative distribution function. Since the scale parameter can be eliminated, the parameter determination process requires the iterative solution on a personal computer (PC) of only two equations. Convergence of the iteration process provides the ultimate practical justification for the applicability of the considered distribution class relative to given empirical data. Examples are given to verify the efficiency of the proposed parameter determination method.

I. INTRODUCTION

The objective is to revitalize interest in the application of a class of probability distributions which had been designated a generalized Gamma distribution by various authors [1, 2, 3, 4, and 5]. This class represents three-parameter, continuous, one-sided distributions which may be defined in terms of the cumulative distribution function (cdf)

$$F(x) = \begin{cases} \frac{1}{\Gamma((1-p)\beta^{-1})} \gamma((1-p)\beta^{-1}, \varepsilon\beta), & \varepsilon = xb^{-1}, x > 0, \\ 0, & x \leq 0, \end{cases} \quad (*)$$

with parameters b , p , and β , $\Gamma(y)$ and $\gamma(a, y)$ being the Gamma function and the incomplete Gamma function (with lower integration limit zero), respectively.

Apparently, this class of distributions was introduced originally by L. Amoroso [6]. Various aspects of it received attention in fairly recent publications [7, 8]. These papers refer to the close connection of the class (*), via the associated probability density function (pdf), with a class of parabolic differential equations (generalized Feller equation). They also establish a connection with the underlying dynamical diffusion process. In this context the publication [9, Sec. 7] may be of particular interest.

The probability density function (pdf) class associated with the cdf class (*) is given by

$$f(x) = \frac{dF(x)}{dx} = \begin{cases} \frac{\beta}{\Gamma((1-p)\beta^{-1})} b^{-1} \varepsilon^{-p} \exp -\varepsilon\beta, & \varepsilon = xb^{-1}, x > 0, \\ 0, & x \leq 0. \end{cases} \quad (1)$$

The expression for $f(x)$ clearly demonstrates the meaning of the parameters b , p , and β . The parameter $b > 0$ represents scale, $p < 1$ represents initial shape (for small values of $x > 0$) and $\beta > 0$ represents terminal shape (for large values of x).

A shift parameter x_0 may be introduced by replacing x by $x-x_0$, $x > x_0$. That will not be done here, however, since only distributions of the three-parameter type (*), (1), are of interest. To partially lift the restrictions on p and β , one may replace the independent variable x by, say, y^{-1} , $y > 0$ being a new independent variable; however, this possibility will not be of further concern here. Another remark concerns a notational change relative to the earlier papers [7, 8, 9]. The parameter λ which appeared there has been replaced by $\beta = 1 - \lambda$.

The reason for the designation of p and β as initial and terminal shape parameters, respectively, is evident. For large values of x , the exponential function in (1) is the dominating factor and, consequently, the shape of the pdf curve or, more precisely, its rate of decay, for large values of x is determined by β . In any case, $f(x) \rightarrow 0$ as $x \rightarrow +\infty$. Since the exponential function approaches unity as $x \rightarrow 0$, the initial shape of the pdf curve is determined by p . If $0 < p < 1$, $f(x) \rightarrow +\infty$ as $x \rightarrow 0$ so that, in this case, a J-shaped distribution is being dealt with. If $p = 0$, $f(x) \rightarrow \beta / b\Gamma(b\Gamma(\beta^{-1}))$; the distribution is of the half bell-shaped type (purely exponential). Finally, if $p < 0$, $f(x) \rightarrow 0$ as $x \rightarrow 0$. The distribution is hump-shaped, the pdf having a unique maximum at the point $x_m = b(-p\beta^{-1})^{1/\beta}$.

For particular values of the shape parameters, the class of distributions characterized by (*) contains a number of special cases well-known in statistics and statistical physics. The major ones are: [2, 7, 10]:

Gauss	$(p = 0, \beta = 2),$
Weibull	$(p = 1 - \beta < 1),$
exponential	$(p = 1 - \beta = 0),$
Rayleigh	$(p = 1 - \beta = -1),$
Gamma	$(p < 1, \beta = 1),$
chi-square	$(p = (2 - \nu)/2 < 1, \beta = 1),$
Maxwell	$(p = -2, \beta = 2, x = vt_0, b = (2kT/m)^{1/2}t_0),$ and
Wien	$(p = -3, \beta = 1, x = 2\pi c\omega_0^{-2}\omega, b = 2\pi c\omega_0^{-2}kTh^{-1}.$

Apparently, application of the distribution class (*) has been severely limited, although various attempts have been made, [1, 2, 3, 4, and 11] for the special cases of Gamma and Weibull to formalize and standardize the parameter estimation process. In fact, the distribution class (*), has not been used as extensively in every day statistical practice as it should be. The main reason for this state of affairs is most likely attributed to computational intensity and possibly to convergence problems arising in the numerical solution of the associated maximum-likelihood equations. This report will not deal further with questions related to the maximum-likelihood approach. This will be done elsewhere in a separate publication.

From an application point of view, to revitalize interest in the distribution class (*) means to provide a practically useful, efficient, and computationally economical technique for the determination of the three unknown parameters b , p , and β relative to given frequency data. Practical usefulness implies the notion of a criterion being involved whose satisfaction can be verified in the application of the technique. The parameter determination technique that is being proposed here does involve such a criterion. It is based on an applicability criterion, announced already [7], which is characteristic for the distribution class (*). This criterion, which will be presented in Section II, recognizes the fact that the logarithm of the cdf, is asymptotically linear in $\log x$ as x approaches zero from above. This typical property of the class (*), can be exploited to establish one equation in the three unknowns b , p , and β , which encompasses the cumulative frequency data. Two more equations in the three unknowns are obtained from the first and second moments which can be numerically determined from the relative frequency data. Since the scale parameter b can easily be eliminated by means of the first moment, two equations are eventually left in the unknowns p and β .

The equation resulting from the log cdf function is too complicated to be used directly. Therefore, it will be replaced by a simpler approximating function which will subsequently be used for a least squares fit of the given log cdf points. The quality of this approximation will be discussed in Appendixes B and C.

The solution of the two final equations for the two unknown parameters p and R proceeds by iteration (Section V). Convergence of the iteration process provides the ultimate practical justification for the application of the distribution class (*), relative to given empirical data (Appendix D).

A number of examples are presented in Section VI. These are "synthetic" examples in the sense that their parameter values are known in advance and then reconstructed by means of the proposed parameter determination method. A quality test is immediately available by means of comparison of the original and the calculated parameter values. One empirical example has been included for purposes of exposition and demonstration. No attempt will be made in this report to do a goodness-of-fit test. This will be left to another publication which will deal exclusively with empirical examples.

While work on this project was in progress and during its publication phase, parallel efforts on maximum-likelihood density estimation for the hyper-Gamma class have led to essential new results [14] which cover both the three- and four-parameter cases. Although computer programming via the maximum-likelihood approach is more complex than that required by the technique presented in this report, maximum-likelihood density estimation may be preferable in practice. Nevertheless, the method presented here leads quickly to approximate parameter values which may be used as initial values in maximum-likelihood estimations.

II. NOTATIONS AND FORMULAS

In statistical practice, empirical data are normally given in terms of absolute frequencies, f_a , relative to a finite number m of class intervals, $[x_{v-1}, x_v)$ ($v = 1, \dots, m$). The intervals are assumed to be of equal length, $d = x_v - x_{v-1}$, so that $x_v = vd$, and $x_0 = 0$.

The (piecewise constant) absolute frequency function $f_a(x)$, $x \in [0, x_m)$, is defined as $f_a(x) = f_a(x_{v-1})$ for $x \in [x_{v-1}, x_v)$. A relative frequency function, $f_r(x)$, $x \in [0, x_m)$ can now be defined as $f_r(x) = N^{-1}f_a(x)$, N being the total number of observations, i.e.,

$$N = \sum_{v=1}^m f_a(x_v).$$

With $f_r(x)$ one associates the (empirical) pdf $f(x) = d^{-1}f_r(x)$, $x \in [0, x_m)$. A major problem in statistical analysis arises in the attempt to construct a continuous analogue of a given (piecewise constant) empirical pdf. The main objective of the work to be presented in this report deals with a new approach to the solution of this problem within the class of distributions (*).

The (empirical) cdf associated with given frequency data is defined as a continuous and piecewise linear function $F(x)$ with functional values at $x = 0$ and at the interval endpoints given by

$$F(0) = 0, \\ F(x_v) = \sum_{u=1}^v f(x_{u-1})d = \sum_{u=1}^v f_r(x_{u-1})(v = 1, \dots, m). \quad (2)$$

The set of m class intervals, $[x_{v-1}, x_v)$, $x_0 = 0$, $x_v = vd$ ($v = 1, \dots, m$) together with the cdf values, $F(x_v)$, as defined in (2), shall be called an empirical data set.

The (theoretical) moments of the distribution class (*) are given by the formula

$$M_v = \int_0^{\infty} x^v f(x) dx = b^v \frac{\Gamma((v+1-p)R^{-1})}{\Gamma((1-p)R^{-1})} \quad (v = 0, 1, 2, \dots), \quad (3)$$

$f(x)$ given by (1), $M_0 = 1$, $M_1 = \mu$ being the mean value, and M_2 being the mean square value.

Observe the important inequality

$$0 < \frac{\mu^2}{M_2} < 1 \quad (4)$$

which follows from

$$0 < \int_0^{\infty} (x-\mu)^2 f(x) dx = M_2 - \mu^2.$$

Replacement of $f(x)$ in (3) by the empirical pdf yields the (empirical) first and second moments,

$$M_1 = d \sum_{v=1}^m f_r(x_v)(v - 1/2), \quad (5)$$

$$M_2 = d^2 \sum_{v=1}^m f_r(x_v)(v(v-1) + 1/3). \quad (6)$$

III. THE APPLICABILITY CRITERION

Now return to the cdf class, $F(x)$, given in (*). By means of the definition of the incomplete Gamma function $\gamma(a, y)$ in terms of the degenerate hyper-geometric function $\phi(.,.,.)$ [12; 9.236.4] the nontrivial part of $F(x)$ can be represented in the form

$$F(x) = \frac{1}{\Gamma(1+(1-p)\beta^{-1})} \xi^{1-p\phi((1-p)\beta^{-1}, 1 + (1-p)\beta^{-1}; -\xi^\beta)},$$

$$\xi = xb^{-1}, x > 0.$$

This allows a useful expression for the logarithm of $F(x)$ to be obtained:

$$\begin{aligned} \log F(x) = & -\log \Gamma(1+(1-p)\beta^{-1}) + (1-p)\log xb^{-1} \\ & + \log \phi((1-p)\beta^{-1}, 1 + (1-p)\beta^{-1}; - (xb^{-1})^\beta). \end{aligned} \quad (7)$$

The independent variable transformation $x = M_1 y$ is carried out. The reason for this transformation is that, for a given empirical data set, all interval endpoints $x_v = v d$ with $x_v < M_1$ will be transformed into points y_v with $0 < y_v < 1$, so that the corresponding numbers $u_v = \log y_v = \log x M_1^{-1}$ will be negative. (In some cases where there are only a few points $x_v < M_1$, it may be better to transform x into y by means of a factor κ , $M_1 < \kappa \leq x_m$. In any case, from a practical point of view, as will be seen shortly, it is essential to have "sufficiently" many numbers $u_v = \log y_v = \log x_v \kappa^{-1}$ with $u_v < 0$.) With $\log y = u$ and $\log F(x) = \log F(M_1 y) = \log F(M_1 e^u) = v(u)$, so that $\log x b^{-1} = \log y M_1 b^{-1} = u - \log M_1^{-1} b$, the functional relation

$$\begin{aligned} v(u) = & (1-p)u - \log \Gamma(1+(1-p)\beta^{-1}) - (1-p) \log M_1^{-1} b \\ & + \log \phi((1-p)\beta^{-1}, 1 + (1-p)\beta^{-1}; - (M_1 b^{-1} e^u)^\beta) \end{aligned} \quad (8)$$

is obtained from (7). The function ϕ is represented as a power series in its last argument with constant term equal to unity. Therefore, as $x \downarrow 0$, i.e., as $y \downarrow 0$, which means as $u \downarrow -\infty$, $\log \phi \downarrow 0$. (For the argument of ϕ in (8) the series is alternating and, hence, $0 < \phi < 1$.) Consequently, the function $v(u)$ given in (8) is asymptotically linear in u as $u \downarrow -\infty$. In other words,

$$v(u) \sim v_a(u) = (1-p)u - \log \Gamma(1+(1-p)\beta^{-1}) - (1-p) \log M_1^{-1} b, \quad u \downarrow -\infty.$$

This asymptotic linearity property may also be expressed by saying that, as $u \downarrow -\infty$, the graph of the function $v(u)$ approaches the (straight line) asymptote determined by the equation

$$v_a(u) = (1-p)u - \log \Gamma(1+(1-p)\beta^{-1}) - (1-p) \log M_1^{-1} b.$$

Here and in (8) the scale parameter b may be eliminated by means of the first moment, using (3) for $v = 1$,

$$b = M_1 \frac{\Gamma((1-p)\theta^{-1})}{\Gamma((2-p)\theta^{-1})},$$

which leads to

$$v_a(u) = (1-p)u - \log(1-p)\theta^{-1} - (2-p)\log\Gamma((1-p)\theta^{-1}) + (1-p)\log\Gamma((2-p)\theta^{-1}). \quad (9)$$

Obviously, the graph of the function $v(u)$ has a second asymptote, namely the line $v = 0$ for $u \rightarrow \infty$. This one, however, is of no further interest.

Based on the asymptotic linearity property of the function $v(u)$, one can formulate the following applicability criterion which has been announced already in [7]:

A distribution function $F(x)$ of the class (*) may be considered as a candidate for a data fit if the logarithmic plot of a given set of empirical data, i.e., the plot of the points $P_v = (u_v, v_v)$, $u_v = \log x_v \kappa^{-1}$, $M_1 < \kappa < x_m$, $v_v = \log F(x_v)$ ($v = 1, \dots, m$), indicates the existence of an asymptote as $u \rightarrow \infty$.

It is essential to observe that the initial shape parameter p of a member of the distribution class (*) is uniquely determined by the direction angle θ of the asymptote of the graph of the function $v(u)$. According to (8) and (9), $1-p = \tan \theta$. This fact will be exploited in the parameter determination method.

IV. DETERMINATION OF THE PARAMETERS

This section presents the general outline of the proposed parameter determination method relative to the distribution class (*). The actual computational procedure will be established in Section V.

Determination of the parameters b , p , and β relative to a given empirical data set requires the solution of three simultaneous equations. For notational convenience p is replaced by $1-\sigma$. Since the scale parameter b can be expressed uniquely in terms of the two shape parameters by means of the first moment (3), it is actually necessary to have only two equations involving the two shape parameters. One such equation can be obtained from the second moment upon elimination of b . It is of the form

$$h(\beta, \sigma) = \Gamma^2 \left(\frac{1+\sigma}{\beta} \right) - A \Gamma \left(\frac{\sigma}{\beta} \right) \Gamma \left(\frac{2+\sigma}{\beta} \right) = 0, A = \frac{M_1^2}{M_2}, \quad (10)$$

in which, according to (4), $0 < A < 1$. A second equation, $g(\beta, \sigma) = 0$, follows from the function $v(u)$ given in (8) if $u = 0$ and b is eliminated.

Unfortunately, the second equation is unpleasant from a computational point of view. It is desirable, therefore, from a practical standpoint, to replace it by some other equation which can more easily be handled.

To achieve this objective, an approximating function $v^*(u)$ is used for the function $v(u)$ with the fact in mind that the asymptote of the graph of $v(u)$ determines the initial shape parameter uniquely. For $v^*(u)$ the function

$$v^*(u) = \sigma u + \rho(e^{\beta u} - 1) + v(0), \quad \sigma = 1-p \quad (11)$$

is chosen.

There are several reasons for this choice of $v^*(u)$:

(1) The graph of $v^*(u)$ has the asymptote $v_a^*(u) = \sigma u - \rho + v(0)$ as $u \rightarrow -\infty$, its direction tangent $\sigma = 1-p$ being the same as that of the asymptote of the graph of the original function $v(u)$ (9),

(2) The function $v^*(u)$ approximates the function $v(u)$ well over the interval $(-\infty, 0]$ (Appendix B). Of course, regardless of the value ρ , $v^*(u)$ will not approximate $v(u)$ for large values of u , since $v(u) \rightarrow 0$ as $u \rightarrow +\infty$ whereas $v^*(u)$ does not. This is no matter of concern, however. The intention is to exploit the asymptotic linearity property of $v(u)$ as $u \rightarrow -\infty$.

(3) $v^*(0) = v(0)$, and

(4) The function $v^*(u)$ is linear in its coefficients σ and ρ .

If $v(u)$ can now be approximated by $v^*(u)$ in such a fashion that the coefficient σ , say, becomes a well-defined function of β , $\sigma = \sigma(\beta)$, then the needed second equation, $g^*(\beta, \sigma) = \sigma - \sigma(\beta) = 0$ to solve the problem results.

The easiest way to explain the procedure is to go along with an example. Table 1 shows absolute frequencies f_a (FABS) over $m=14$ classes (K) with intervals $(x_{v-1}, x_v) = [v-1, v)$ of length $d = 1$. The total number of observations is $N = 119$. The data for this example (Example Library Classification: EMPEX #3) originated from Reference [13]. EMPEX #3 presents the frequency distribution of 119 upper-tropospheric wind speeds measured over Nashville, Tennessee between mid-May and mid-September 1985. The reported (scalar) wind speed values refer to the 300 hektopascal level which corresponds approximately to a height of 9.6 km. The original reports [13] of wind speeds in integral values of knots have been grouped here into classes of 5 knots. Therefore, the v th class interval $[v-1, v)$ contains the observations from $5v-5$ to $5v-1$ knots ($v=1, \dots, 14$).

TABLE 1. Empirical Example #3 - Absolute Frequencies.

K	XR	FABS
1:	1.00	2 *****
2:	2.00	6 *****
3:	3.00	14 *****
4:	4.00	17 *****
5:	5.00	21 *****
6:	6.00	14 *****
7:	7.00	15 *****
8:	8.00	10 *****
9:	9.00	7 *****
10:	10.00	6 *****
11:	11.00	2 *****
12:	12.00	1 ***
13:	13.00	3 *****
14:	14.00	1 ***

The relative frequencies f_r (FREL) are given in Table 2 together with the cdf values F (CUMREL) at the right-hand interval endpoints (XR) calculated according to (2). This table also shows the coordinates $u_v = \log vM_1^{-1}$, $v_v = \log F(v)$ of the log cdf points $P_v = (u_v, v_v)$ in the U- and V- columns. The value of $M_1 = 5.4496$ has been determined from (5). (It corresponds to 26.248 knots).

TABLE 2. Empirical Example #3 - Relative Frequencies.

K	XR	FREL	CUMREL	U	V
1	1.00	1.68%	1.68%	-1.6955	-4.0860
2	2.00	5.04%	6.72%	-1.0024	-2.6997
3	3.00	11.76%	18.49%	-0.5969	-1.6881
4	4.00	14.29%	32.77%	-0.3092	-1.1156
5	5.00	17.65%	50.42%	-0.0861	-0.6848
6	6.00	11.76%	62.18%	0.0962	-0.4751
7	7.00	12.61%	74.79%	0.2504	-0.2905
8	8.00	8.40%	83.19%	0.3839	-0.1840
9	9.00	5.88%	89.08%	0.5017	-0.1157
10	10.00	5.04%	94.12%	0.6070	-0.0606
11	11.00	1.68%	95.80%	0.7024	-0.0429
12	12.00	0.84%	96.64%	0.7894	-0.0342
13	13.00	2.52%	99.16%	0.8694	-0.0084
14	14.00	0.84%	100.00%	0.9435	0.0000

The plot of the points P_v (with $P_9, P_{10}, P_{12}, P_{13}$ omitted for reasons of clarity) is shown in Figure 1. Inspection of the plot leads to the conclusion that the class (*) can be applied for a data fit.

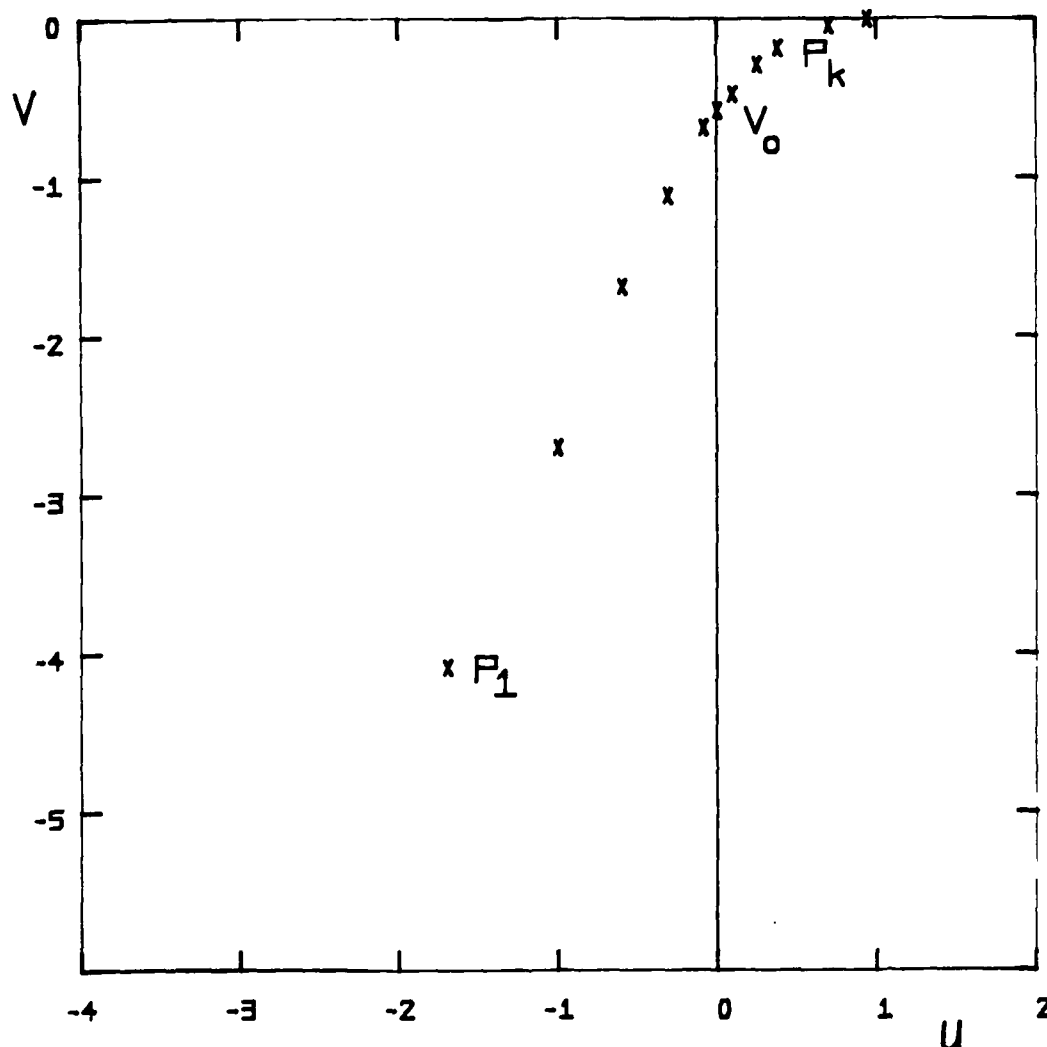


Figure 1. Plot of points P_v .

Digressing briefly, a few remarks concerning cdf plots like the one shown in Figure 1 are offered. It is strongly recommended that the plot be prepared for a given empirical data set and inspected carefully for the following reasons: (1) It provides the first opportunity to decide whether or not the distribution class (*) should be applied for a data fit, and (2) the plot provides the analyst with some basic information about the type of distribution he is dealing with beyond that which can be extracted from a histogram. If an asymptote location can be estimated, its direction angle θ provides immediately an estimate of the initial shape parameter p since $\tan \theta = \sigma = 1-p$. Observe that $0 < p < 1$ (J-shaped pdf) if $0 < \theta < \pi/4$, $p = 0$ if $\theta = \pi/4$ (half bell-shaped type pdf, purely exponential), and $p < 0$ (hump-shaped pdf) if $\pi/4 < \theta < \pi/2$.

If now, in the general case, the plot of points $P_v = (u_v, v_v) (v=1, \dots, m)$, $u_v = \log x_v M_1^{-1}$, $v_v = \log F(x)$, with enumeration done such that $u_1 < u_2 < \dots < u_{\kappa-3} < 0 < u_{\kappa-2} < \dots < u_m$, of a given empirical data set indicates that the distribution class (*) is applicable, then there must be numbers $p = 1 - \sigma$ and β (and b) such that the function $v(u)$ "fits" the points P_v . If this is so, then if $v^*(u)$ is a good approximating function of $v(u)$, the same will be true for $v^*(u)$ if the parameters σ , ρ , and β have been properly chosen.

To specify the coefficients σ and ρ of $v^*(u)$, perform a least squares fit on the points $P_v = (u_v, v_v)$ with

$$u_1 < u_2 < \dots < u_{\kappa-3} < 0 < u_{\kappa-2} < u_{\kappa-1} < u_{\kappa}, \quad (12)$$

disregarding all others with index greater than κ . In the above example, $\kappa = 8$. The reasons for this choice of a subset of the points P_v are that, (1) points with $u_v < 0$ over which the quality of the fit may be poor are eliminated and (2) a sufficient number of points are available to adequately account for the typical concavity of the graph of the function $v(u)$ (Fig. 1). Experience shows that a minimum of five points P_v with negative abscissas u_v are normally adequate. Should there be less than five such points under scaling of x by means of M_1 , one should use a scaling factor $\kappa < M_1$.

With β in (11) as a parameter, the least squares fit on the points P_v with abscissas (12) leads to a system of two linear equations for σ and ρ which can easily be solved to give σ and ρ as functions of β , $\sigma = \sigma(\beta)$, $\rho = \rho(\beta)$. Actually, only $\sigma = \sigma(\beta)$ is needed for the parameter determination procedure. The function $\rho(\beta)$ is useful, however, to judge the quality of the approximation of $v(u)$ by $v^*(u)$ (Appendix B).

Of course, it is necessary in this process to determine the numerical value of $v(0)$ which appears in (11). But this number can easily be calculated by means of Lagrange-Aitken interpolation over the consecutive points $P_{\kappa-4}$, $P_{\kappa-3}$, $P_{\kappa-2}$, $P_{\kappa-1}$ with $u_{\kappa-4} < u_{\kappa-3} < 0 < u_{\kappa-2} < u_{\kappa-1}$.

V. THE ITERATION PROCESS

The least squares fit on the log cdf points $P_v = (u_v, v_v)$ with abscissas satisfying the inequalities (12) by means of the function $v^*(u)$ given in (14) leads to the error equations

$$v^*(u_v) = \sigma u_v + \rho(e^{\beta u_v} - 1) + v(0) = v_v + \varepsilon_v \quad (v = 1, \dots, \kappa). \quad (13)$$

Minimization of the sum of squares of the errors ε_v specifies the coefficients σ and ρ as functions of the parameter β ,

$$\sigma(\beta) = D_1/D, \quad \rho(\beta) = D_2/D. \quad (14)$$

The determinants are defined by

$$D(\beta) = A_{11}A_{22} - A_{12}^2, \quad D_1(\beta) = BA_{22} - CA_{12}, \quad D_2(\beta) = CA_{11} - BA_{12} \quad (15)$$

with

$$\begin{aligned} A_{11} &= \sum_{v=1}^{\kappa} u_v^2, \quad A_{12}(\beta) = \sum_{v=1}^{\kappa} u_v a_v, \quad A_{22}(\beta) = \sum_{v=1}^{\kappa} a_v^2, \\ B &= \sum_{v=1}^{\kappa} u_v c_v, \quad C(\beta) = \sum_{v=1}^{\kappa} a_v c_v, \quad a_v = e^{\beta u_v} - 1, \quad c_v = v_v - v(0). \end{aligned} \quad (16)$$

In addition to the equation $g^*(\beta, \sigma) = \sigma - \sigma(\beta) = 0$, use the equation $h(\beta, \sigma) = 0$ given in (10). The coefficient A which appears in the function h is to be determined by means of the formulas (5) and (6). Essential properties of the equations $g^* = 0$ and $h = 0$ are discussed in Appendixes A and C.

The iteration process now proceeds as follows. Set $\beta = 1$ in (13) and calculate the value $\sigma_1 = \sigma(1)$ from (14). Then solve the equation $h(\beta, \sigma_1) = 0$. As a matter of fact, use of the equation $H(\alpha, \sigma) = 0$ obtained from $h(\beta, \sigma) = 0$ by the substitution $\beta = \alpha(1-\alpha)^{-1}$ reduces the interval of the unknown from $(0, +\infty)$ to $(0, 1)$. The regula falsi method is used with the starting value $\alpha = 0.5$, and a search for the first pair of functional values of opposite sign is initiated. Iteration is terminated when $|\alpha_v - \alpha_{v-1}| < 10^{-3}$. (The full Newton's method has also been used with no essential improvement in accuracy but the added computational burden of having to evaluate the psi function.)

The solution β_1 of $h(\beta, \sigma_1) = 0$ is then used to calculate the value $\sigma_2 = \sigma(\beta_1)$ from (14). Proceeding in this fashion, establish two sequences $\{\sigma_v\}$ and $\{\beta_v\}$ which, provided the data set is well-conditioned, will converge (Appendix D) to numbers $\sigma_0 = 1 - p_0$ and β_0 , respectively. These numbers p_0 and β_0 are the final values for the shape parameters p and β . The final value b_0 for the scale parameter b is then obtained from the first moment, $b_0 = M_1 \Gamma((1-p_0)\beta_0^{-1}) / \Gamma((2-p_0)\beta_0^{-1})$, and the parameter determination process is complete.

In practice, of course, the iteration process will be terminated when a desired accuracy has been reached. For the examples to be discussed in the next section, the criterion $|\sigma_v - \sigma_{v-1}| < 10^{-2}$, $|\beta_v - \beta_{v-1}| < 10^{-2}$ was used and seems to be adequate. Thus, the first pair of values σ_v and β_v which satisfy this criterion was taken as the final values.

VI. EXAMPLES

This section presents a number of examples for the parameter determination method. To demonstrate its efficiency, several special cases of the distribution class (*) were selected for which, in order to be able to evaluate the results objectively, the parameter values were chosen to begin with and then reconstructed. The resulting errors in these examples are entirely due to errors arising from the approximation of the function $v(u)$ by the function $v^*(u)$. The $\log F(x)$ values have been calculated directly from the exact cdf's which, since all of the examples are of Weibull type (i.e., $p = 1 - v$), are given by

$$F(x) = 1 - \exp - \xi^{\beta}, \quad \xi = xb^{-1}. \quad (17)$$

The moments M_1 and M_2 have been calculated from the formula (3) by means of the given b , p , and β values. In empirical cases, additional errors will arise from the use of the sample moments.

There are four examples, classified in our example library as SYNEX (= synthetic example) #8, #9, #10, and #11. SYNEX #8 represents an exponential distribution, SYNEX #11 a J-shaped distribution. The others are of hump-shaped type. SYNEX #10 is being presented in two different versions relative to the number of classes.

Tables 3, 4, 5, 6, and 7 are essentially self-explanatory. The heading includes the original parameter values ($p = 1 - \beta$ in all cases). Column K indicates the class interval number. In the second column, $x_v = XR$ gives the right-hand class interval endpoint. The interval length d in each case can immediately be extracted from this column. The cdf values $F(x_v) = \text{CUMREL}$ are shown in column 4 as calculated from (17) for the given b , p , and β values up to values of x_v in such a way that the first three points $P_v = (u_v, v_v)$ ($v = \kappa - 2, \kappa - 1, \kappa$) in the fourth quadrant of the (u, v) -plane are included in the set of points to be used for the least squares fit. Column 3 (which is actually of no interest relative to the SYNEX's) shows the relative frequencies $f_r(x_v) = \text{FREL}$ calculated from the cdf values. The coordinates $u_v = U$, $v_v = V$ of the points P_v are given in columns 5 and 6, respectively. The last column $DV/DU = \tan \theta_{m-v-1}$ contains the coordinate difference ratios. It is of some interest in these SYNEX's only.

The moments $M_1 = M1$, $M_2 = M2$, and the numbers $A = M_1^2/M_2$ and $v(0) = V0$ are given in the center block of each table. In each case, the numerical value of $v(0)$ has been calculated by means of four-point-Lagrange-Aitken interpolation as explained at the end of Section IV.

The last block in each table contains the numerical results for each iteration step. The final values for the parameters appear in the lower right-hand corner. Iteration in each example has been started with $\beta = 1$ and terminated at $|\beta_v - \beta_{v-1}| < 10^{-2}$.

It should be observed that the fact that the examples are of Weibull type has nowhere been used in the iteration process, i.e., p and β (and b) have been individually determined.

TABLE 3. SYNEX #8: Exponential Weibull Distribution (B=1, P=0, BETA=1)

K	XR	FREL	CUMREL	U	V	DV/DU
1	0.10	9.52%	9.52%	-2.303	-2.352	1.022
2	0.20	8.61%	18.13%	-1.609	-1.708	0.930
3	0.30	7.79%	25.92%	-1.204	-1.350	0.882
4	0.40	7.05%	32.97%	-0.916	-1.110	0.836
5	0.50	6.38%	39.35%	-0.693	-0.933	0.793
6	0.60	5.77%	45.12%	-0.511	-0.796	0.751
7	0.70	5.22%	50.34%	-0.357	-0.686	0.711
8	0.80	4.73%	55.07%	-0.223	-0.597	0.672
9	0.90	4.28%	59.34%	-0.105	-0.522	0.635
10	1.00	3.87%	63.21%	0.000	-0.459	0.599
11	1.10	3.50%	66.71%	0.095	-0.405	0.566
12	1.20	3.17%	69.88%	0.182	-0.358	0.533
13	1.30	2.87%	72.75%	0.262	-0.318	0.502
M1 = 1.0000 M2 = 2.0000 A = 0.5000 V0 = -0.4587						
Iteration #1: RHO = -0.4093 PO = 0.0191 SIGMA = 0.9809 BETAO= 1.0201 ALPHAO= 0.5050 BO = 1.0479 Iteration #2: RHO = -0.3988 PO = 0.0226 SIGMA = 0.9774 BETAO= 1.0239 ALPHAO= 0.5059 BO = 1.0571						

TABLE 5. SYNEX #10: Weibull Distribution (B=2, P=-2, BETA=3)

K	XR	FREL	CUMREL	U	V	DV/DU
1	0.10	0.01%	0.01%	-2.883	-8.987	3.118
2	0.20	0.09%	0.10%	-2.189	-6.908	2.999
3	0.30	0.24%	0.34%	-1.784	-5.693	2.997
4	0.40	0.46%	0.80%	-1.496	-4.832	2.992
5	0.50	0.75%	1.55%	-1.273	-4.167	2.983
6	0.60	1.11%	2.66%	-1.091	-3.625	2.969
7	0.70	1.53%	4.20%	-0.937	-3.171	2.949
8	0.80	2.00%	6.20%	-0.803	-2.781	2.922
9	0.90	2.51%	8.71%	-0.685	-2.441	2.886
10	1.00	3.04%	11.75%	-0.580	-2.141	2.842
11	1.10	3.58%	15.33%	-0.485	-1.876	2.788
12	1.20	4.10%	19.43%	-0.398	-1.639	2.724
13	1.30	4.59%	24.01%	-0.318	-1.427	2.649
14	1.40	5.02%	29.04%	-0.243	-1.237	2.562
15	1.50	5.38%	34.42%	-0.174	-1.067	2.465
16	1.60	5.65%	40.07%	-0.110	-0.915	2.356
17	1.70	5.82%	45.89%	-0.049	-0.779	2.236
18	1.80	5.87%	51.76%	0.008	-0.659	2.107
19	1.90	5.81%	57.57%	0.062	-0.552	1.968
20	2.00	5.64%	63.21%	0.113	-0.459	1.822

M1 =	1.7860	M2 =	3.6110	A =	0.8833
		V0 =	-0.6746		

Iteration #1:	RHO = -0.7792	PO = -2.1569
	SIGMA = 3.1569	BETA0= 2.7861
	ALPHA0= 0.7359	BO = 1.8925
Iteration #2:	RHO = -0.3418	PO = -2.0029
	SIGMA = 3.0029	BETA0= 2.9976
	ALPHA0= 0.7498	BO = 1.9985
Iteration #3:	RHO = -0.3276	PO = -1.9963
	SIGMA = 2.9963	BETA0= 3.0075
	ALPHA0= 0.7505	BO = 2.0031

TABLE 6. SYNEX #10: Weibull Distribution (B=2, P=12, BETA=3)

K	XR	FREL	CUMREL	U	V	DV/DU
1	0.30	0.34%	0.34%	-1.784	-5.693	3.191
2	0.60	2.33%	2.66%	-1.091	-3.625	2.983
3	0.90	6.05%	8.71%	-0.685	-2.441	2.922
4	1.20	10.72%	19.43%	-0.398	-1.639	2.789
5	1.50	14.99%	34.42%	-0.174	-1.067	2.563
6	1.80	17.34%	51.76%	0.008	-0.659	2.238
7	2.10	16.82%	68.58%	0.162	-0.377	1.825
8	2.40	13.66%	82.24%	0.296	-0.196	1.360
M1 = 1.7860 M2 = 3.6110 A = 0.8833 V0 = -0.6746						
Iteration #1: RHO = -1.3093 PO = -2.4482 SIGMA = 3.4482 BETAO= 2.4435 ALPHAO= 0.7096 BO = 1.6902 Iteration #2: RHO = -0.3908 PO = -2.0325 SIGMA = 3.0325 BETAO= 2.9538 ALPHAO= 0.7471 BO = 1.9778 Iteration #3: RHO = -0.2972 PO = -1.9730 SIGMA = 2.9730 BETAO= 3.0435 ALPHAO= 0.7527 BO = 2.0195 Iteration #4: RHO = -0.2840 PO = -1.9643 SIGMA = 2.9643 BETAO= 3.0574 ALPHAO= 0.7535 BO = 2.0257 Iteration #5: RHO = -0.2821 PO = -1.9630 SIGMA = 2.9630 BETAO= 3.0595 ALPHAO= 0.7537 BO = 2.0266						

TABLE 7. SYNEX #11: Weibull Distribution (B=5, P=0.5, BETA=0.5)

K	XR	FREL	CUMREL	U	V	DV/DU
1	1.00	36.06%	36.06%	-2.303	-1.020	0.443
2	2.00	10.81%	46.87%	-1.609	-0.758	0.378
3	3.00	7.04%	53.91%	-1.204	-0.618	0.345
4	4.00	5.20%	59.12%	-0.916	-0.526	0.320
5	5.00	4.10%	63.21%	-0.693	-0.459	0.300
6	6.00	3.35%	66.56%	-0.511	-0.407	0.283
7	7.00	2.81%	69.37%	-0.357	-0.366	0.268
8	8.00	2.40%	71.77%	-0.223	-0.332	0.255
9	9.00	2.08%	73.86%	-0.105	-0.303	0.243
10	10.00	1.83%	75.69%	0.000	-0.279	0.232
11	11.00	1.62%	77.31%	0.095	-0.257	0.222
12	12.00	1.45%	78.76%	0.182	-0.239	0.213
13	13.00	1.30%	80.06%	0.262	-0.222	0.205

M1 = 10.0000	M2 = 600.0000	A = 0.1667
	V0 = -0.2785	

Iteration #1:	RHO = -0.1775	PO = 0.6110
	SIGMA = 0.3890	BETA0 = 0.5897
	ALPHA0 = 0.3709	BO = 11.3206
Iteration #2:	RHO = -0.3785	PO = 0.5568
	SIGMA = 0.4432	BETA0 = 0.5409
	ALPHA0 = 0.3510	BO = 7.5874
Iteration #3:	RHO = -0.4325	PO = 0.5449
	SIGMA = 0.4551	BETA0 = 0.5315
	ALPHA0 = 0.3470	BO = 6.9509

TABLE 8. Empirical Example #3

K	XR	FREL	CUMREL	U	V
1	1.00	1.68%	1.68%	-1.6955	-4.0860
2	2.00	5.04%	6.72%	-1.0024	-2.6997
3	3.00	11.76%	18.49%	-0.5969	-1.6881
4	4.00	14.29%	32.77%	-0.3092	-1.1156
5	5.00	17.65%	50.42%	-0.0861	-0.6848
6	6.00	11.76%	62.18%	0.0962	-0.4751
7	7.00	12.61%	74.79%	0.2504	-0.2905
8	8.00	8.40%	83.19%	0.3839	-0.1840
9	9.00	5.88%	89.08%	0.5017	-0.1157
10	10.00	5.04%	94.12%	0.6070	-0.0606
11	11.00	1.68%	95.80%	0.7024	-0.0429
12	12.00	0.84%	96.64%	0.7894	-0.0342
13	13.00	2.52%	99.16%	0.8694	-0.0084
14	14.00	0.84%	100.00%	0.9435	0.0000
M1 = 5.4496 M2 = 37.1064 A = 0.8003 V0 = -0.5792					
Iteration #1: RHO = -1.1896 PO = -1.6938 SIGMA = 2.6938 BETA0= 1.5309 ALPHAO= 0.6049 BO = 4.0079 Iteration #2: RHO = -0.6288 PO = -1.4580 SIGMA = 2.4580 BETA0= 1.7002 ALPHAO= 0.6297 BO = 4.7563 Iteration #3: RHO = -0.5378 PO = -1.4118 SIGMA = 2.4118 BETA0= 1.7417 ALPHAO= 0.6353 BO = 4.9226 Iteration #4: RHO = -0.5187 PO = -1.4017 SIGMA = 2.4017 BETA0= 1.7512 ALPHAO= 0.6365 BO = 4.9597					

The accompanying Figures 2, 3, 4, 5, and 6 show the log cdf point plots. In Figure 4 the points P_1 and P_2 are not shown.

The calculations have been performed on an IBM-PC compatible microcomputer (without math co-processor) in compiled MS Basic. For empirical samples of frequency distributions with approximately 40 classes, actual computing time was less than 60 sec.

The evaluation of the differences between the obtained parameter values p_0 , β_0 , b_0 , and the original ones, p , β , b , shows that $\max \{ |p_0 - p|, |\beta_0 - \beta|, |b_0 - b| \}$ is $< 6 \cdot 10^{-2}$ in SYNEX #8, $< 2 \cdot 10^{-2}$ in SYNEX #9, $< 8 \cdot 10^{-3}$ in SYNEX #10a, $< 6 \cdot 10^{-2}$ in SYNEX #10b. In SYNEX #11, $\max \{ |p_0 - p|, |\beta_0 - \beta| \} < 5 \cdot 10^{-2}$, but $1.950 < b_0 - b < 1.951$. The large error in the scale parameter demonstrates the well-known sensitivity of this parameter to small changes in the others for J-shaped distributions. The culprit in this matter, of course, is the error in the initial shape parameter p . Ultimately, this error results from the fact that the class interval length in the example used is too big for this type of distribution.

Before closing this section, briefly return to the empirical example, EMPEX #3 considered in Section IV. The first and second moments are (from (5) and (6), respectively) $M_1 = M1 = 5.4496$ and $M_2 = M2 = 37.1064$ as shown in the center block of Table 8, which also shows the numerical values of $A = M_1^2/M_2$ and $v(0) = V0$. Starting with $\beta = 1$, after the 4th iteration step, the final parameter values b_0 , p_0 , and β_0 shown in the lower right-hand corner of Table 8 are obtained. The least squares approximations include the points F_v for $v = 1, \dots, 8$. No goodness-of-fit test will be performed on the final parameter values in this paper. This will be left to a separate publication which will deal exclusively with empirical examples. However, it is worth mentioning that a recently performed maximum-likelihood estimation of the parameters of EMPEX #3 resulted in the final values $\hat{p} = -1.466$, $\hat{\beta} = 1.720$, and $\hat{b} = 4.795$.

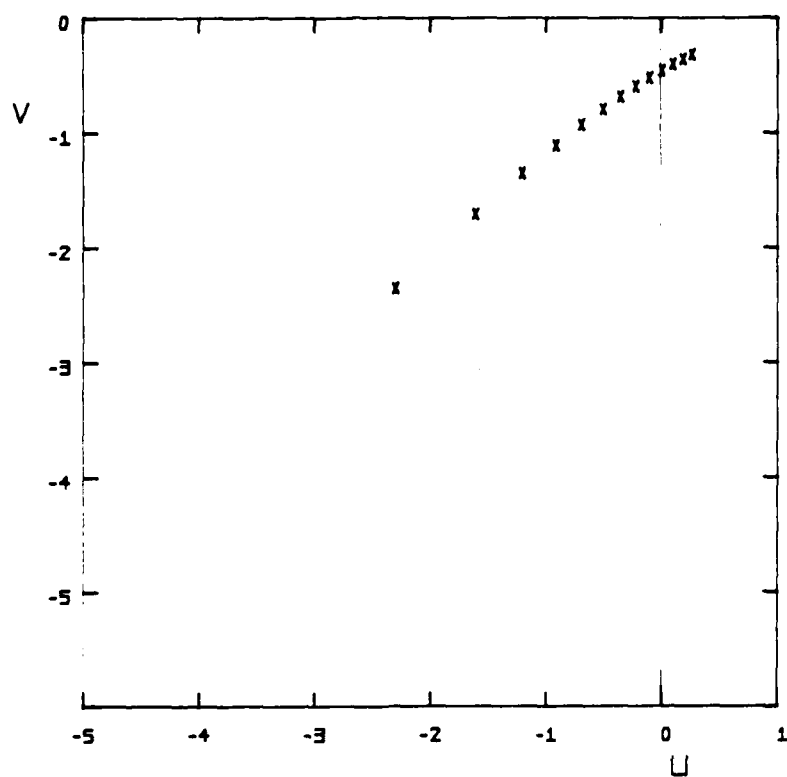


Figure 2

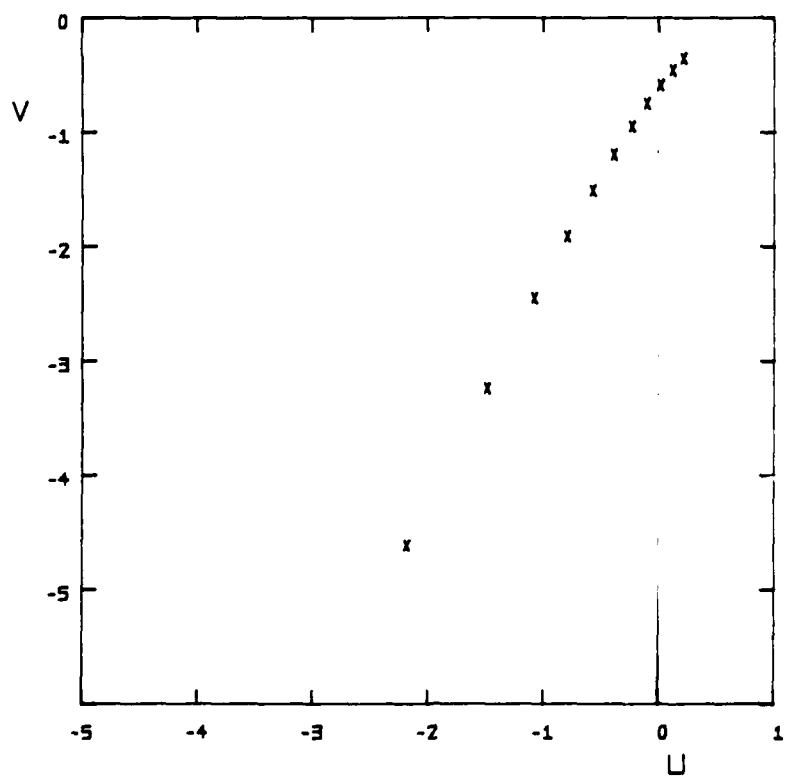


Figure 3

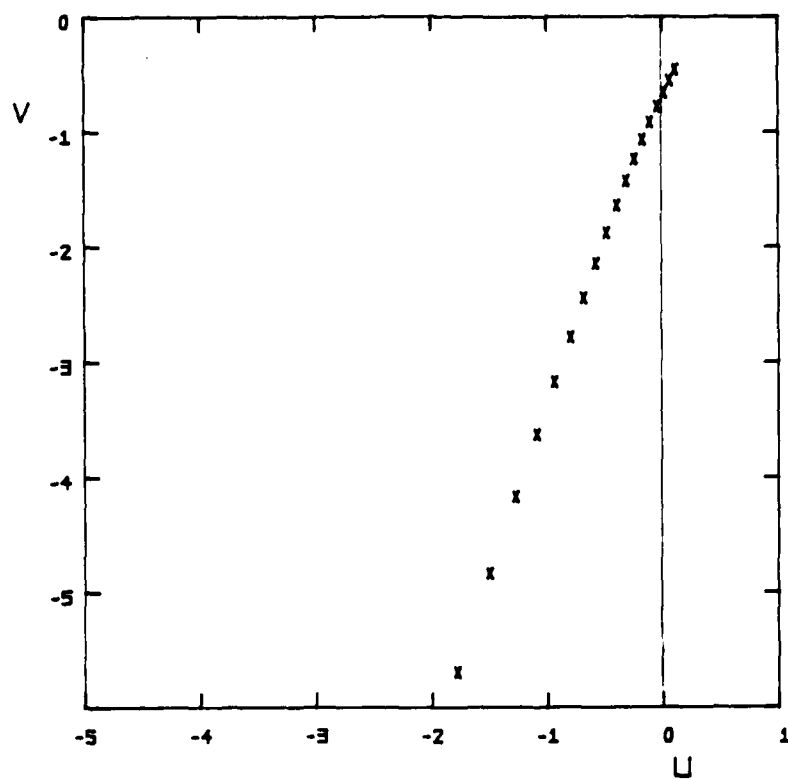


Figure 4

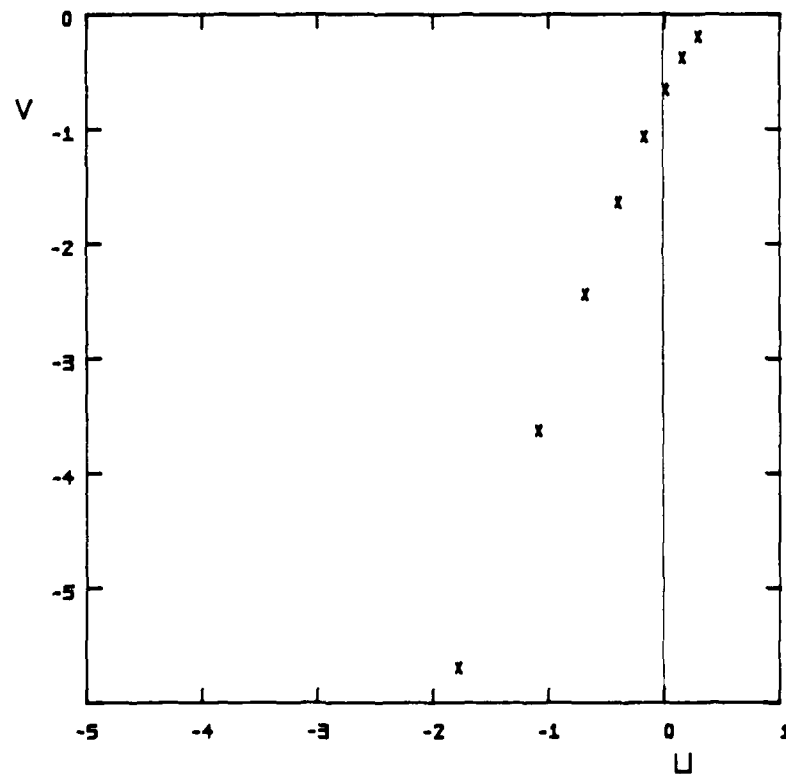


Figure 5

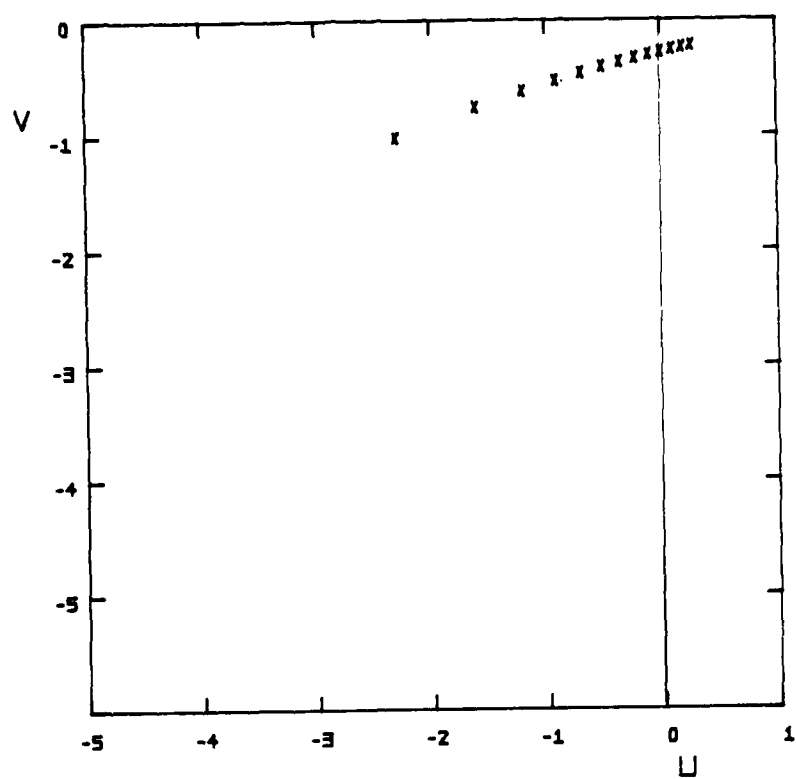


Figure 6

REFERENCES

1. Stacy, E. W.: "A Generalization of the Gamma Distribution," ANN. MATH. STAT. 33, pp. 1187-1192, 1962.
2. Stacy, E. W., and Mihram, G. A.: "Parameter Estimation for a Generalized Gamma Distribution," TECHNOMETRICS 7, pp. 349-358, 1965.
3. Parr, Van B., and Webster, J. T.: "A Method for Discriminating Between Failure Density Functions Used in Reliability Predictions," TECHNOMETRICS 7, pp. 1-10, 1965.
4. Harter, H. L.: "Maximum-Likelihood Estimation of the Parameters of a Four-Parameter Generalized Gamma Population from Complete and Censored Samples," TECHNOMETRICS 9, pp. 159-165, 1967.
5. Essenwanger, O. M.: "APPLIED STATISTICS IN ATMOSPHERIC SCIENCE," Elsevier, Amsterdam, 1976.
6. Amoroso, L.: "Ricerche Intorno Alla Curva Deireditti," ANN. MATH. PARA APPL., ser. 4, 2, pp. 123-157, 1924.
7. Lehnigk, S. H.: "On a Class of Probability Distributions," Accepted by MATH. METH. IN THE APPL. SCI.
8. Lehnigk, S. H.: "Characteristic Functions of a Class of Probability Distributions," Accepted by COMPL. VAR.
9. Lehnigk, S. H.: "Initial Condition Solutions of the Generalized Feller Equation," J. APPL. MATH. PHYS. (ZAMP) 29, pp. 273-294, 1978.
10. Lehnigk, S. H.: "Maxwell and Wien Processes as Special Cases of the Generalized Feller Diffusion Process," J. MATH PHYS. 18, pp. 104-105, 1977.
11. Harter, H. L., and Moore, A. H.: "Maximum-Likelihood Estimation of the Parameters of Gamma and Weibull Populations from Complete and from Censored Samples," TECHNOMETRICS, 7, pp. 639-643, 1965.
12. Gradshteyn, I. S., and Ryzhik, I. M.: TABLES OF INTEGRALS, SERIES, AND PRODUCTS, 4th ed., Academic Press, New York, 1965.
13. NOAA, National Weather Service. Twice daily teletype bulletins of upper-air observations (rawin), Station KBNA (72 327).
14. Lehnigk, S. H., "Maximum-Likelihood Estimation of the Parameters of a Four-Parameter Class of Probability Distributions", Accepted by PROC. EDIN. MATH. SOC.

APPENDIX A

The Equation $h(\beta, \sigma) = 0$

Return to the equation $h(\beta, \sigma) = 0$, $0 < A < 1$, given in (10). The first partial derivatives of $h(\beta, \sigma)$ are

$$h_{\sigma} = \frac{\partial h}{\partial \sigma} = \frac{1}{\beta} \left[2\Gamma^2\left(\frac{1+\sigma}{\beta}\right) \psi\left(\frac{1+\sigma}{\beta}\right) - A\Gamma\left(\frac{\sigma}{\beta}\right)\Gamma\left(\frac{2+\sigma}{\beta}\right) \left(\psi\left(\frac{\sigma}{\beta}\right) + \psi\left(\frac{2+\sigma}{\beta}\right) \right) \right],$$

$$h_{\beta} = \frac{\partial h}{\partial \beta} = - \left[\frac{1}{\beta^2} 2(1+\sigma) \Gamma^2\left(\frac{1+\sigma}{\beta}\right) \psi\left(\frac{1+\sigma}{\beta}\right) - A\Gamma\left(\frac{\sigma}{\beta}\right)\Gamma\left(\frac{2+\sigma}{\beta}\right) \left(\sigma\psi\left(\frac{\sigma}{\beta}\right) + (2+\sigma)\psi\left(\frac{2+\sigma}{\beta}\right) \right) \right],$$

$\psi(y) = d \log \Gamma(y)/dy$ [12;8.360] being the psi function. If $h(\beta^*, \sigma^*) = 0$ for $\beta^* > 0$, $\sigma^* > 0$, then

$$\left(\frac{\partial h}{\partial \sigma} \right)^* = \frac{1}{\beta} \Gamma^2\left(\frac{1+\sigma^*}{\beta^*}\right) 2\psi\left(\frac{1+\sigma^*}{\beta^*}\right) - \psi\left(\frac{\sigma^*}{\beta^*}\right) - \psi\left(\frac{2+\sigma^*}{\beta^*}\right).$$

Use of the series expansion for $\psi(y)$ [12;8.362.1]

$$\psi(y) = -\gamma - \sum_{v=0}^{\infty} \frac{1}{y+v} - \frac{1}{1+v},$$

$\gamma = -\psi(1)$ being Euler's constant [12;8,367.1] results in

$$2\psi\left(\frac{1+\sigma^*}{\beta^*}\right) - \psi\left(\frac{\sigma^*}{\beta^*}\right) - \psi\left(\frac{2+\sigma^*}{\beta^*}\right) = \beta^* \sum_{v=0}^{\infty} r_v,$$

$$r_v = \frac{\sigma^*}{(\sigma^* + v\beta^*)(1+\sigma^* + v\beta^*)(2+\sigma^* + v\beta^*)} > 0, \beta^* > 0, \sigma^* > 0.$$

Consequently,

$$\left(\frac{\partial h}{\partial \sigma} \right)^* = \Gamma^2\left(\frac{1+\sigma^*}{\beta^*}\right) \sum_{v=0}^{\infty} r_v > 0.$$

In other words, considered as a function of $\sigma > 0$ for fixed $\beta > 0$, $h(\beta, \sigma)$ has a positive derivative at each of its zeros σ . Since $h(\beta, \sigma)$ is continuous as a function of σ , it follows that, for fixed $\beta > 0$, $h(\beta, \sigma) = 0$ can have at most one root $\sigma > 0$ and that $h(\beta, \sigma)$ increases from negative to positive values across the root.

Again, if β^* and σ^* are positive numbers such that $h(\beta^*, \sigma^*) = 0$, then

$$\left(\frac{\partial h}{\partial \beta}\right)^* = -\frac{1}{\beta^{*2}} \Gamma^2\left(\frac{1+\sigma^*}{\beta^*}\right) \left[2(1+\sigma^*)\Psi\left(\frac{1+\sigma^*}{\beta^*}\right) - \sigma^*\Psi\left(\frac{\sigma^*}{\beta^*}\right) - (2+\sigma^*)\Psi\left(\frac{2+\sigma^*}{\beta^*}\right) \right].$$

The following formula can be established,

$$2(1+\sigma^*)\Psi\left(\frac{1+\sigma^*}{\beta^*}\right) - \sigma^*\Psi\left(\frac{\sigma^*}{\beta^*}\right) - (2+\sigma^*)\Psi\left(\frac{2+\sigma^*}{\beta^*}\right) = \beta^* \sum_{v=0}^{\infty} s_v,$$

$$s_v = -2\beta^* \frac{v}{(\sigma^* + v\beta^*)(1+\sigma^* + v\beta^*)(2+\sigma^* + v\beta^*)} < 0, \beta^* > 0, \sigma^* > 0,$$

so that

$$\left(\frac{\partial h}{\partial \beta}\right)^* = 2\Gamma^2\left(\frac{1+\sigma^*}{\beta^*}\right) \sum_{v=0}^{\infty} s_v > 0.$$

Therefore, considered as a function of $\beta > 0$ for fixed $\sigma > 0$, $h(\beta, \sigma)$ has a positive derivative at each of its zeros. Continuity again implies that, for fixed $\sigma > 0$, $h(\beta, \sigma) = 0$ has at most one root, $\beta > 0$, and $h(\beta, \sigma)$ increases from negative to positive values across the root.

To continue the investigation of the properties of the equation $h(\beta, \sigma) = 0$ consider the ratio

$$C(\beta, \sigma) = \frac{\Gamma^2\left(\frac{1+\sigma}{\beta}\right)}{\Gamma\left(\frac{\sigma}{\beta}\right) \Gamma\left(\frac{2+\sigma}{\beta}\right)}, \quad (\text{A-1})$$

which is a continuous function of β and σ in the open domain $\beta > 0$, $\sigma > 0$. Set $(1+\sigma)\beta^{-1} = \alpha$, $\beta^{-1} = \gamma$. Then (A-9) can be expressed as an infinite product [12;8.325.1],

$$C(R, \sigma) = \frac{\Gamma(\alpha) \Gamma(\alpha)}{\Gamma(\alpha-\gamma) \Gamma(\alpha+\gamma)} = \prod_{v=0}^{\infty} \left[\left(1 + \frac{1}{(\alpha+v)R}\right) \left(1 - \frac{1}{(\alpha+v)R}\right) \right] = \prod_{v=0}^{\infty} t_v,$$

$$0 < t_v < \frac{(1+\sigma+v\beta)^2-1}{(1+\sigma+v\beta)^2} = 1 - \frac{1}{(1+\sigma+v\beta)^2} < 1, \beta > 0, \sigma > 0. \quad (A-2)$$

The infinite product converges (absolutely) for every $R > 0$ and $\sigma > 0$ since the series

$$\sum_{v=0}^{\infty} (1+\sigma+v\beta)^{-2} < 1+\beta^{-2} \sum_{v=1}^{\infty} v^{-2} < 1 + 2\beta^{-2}.$$

The inequality $0 < t_v < 1$ implies that $0 < C(R, \sigma) < 1$. Furthermore, because of convergence of the product, the series

$$\sum_{v=0}^{\infty} \log t_v = \sum_{v=0}^{\infty} \log \left(1 - \frac{1}{(1+\sigma+v\beta)^2} \right)$$

converges. Its μ th partial sum is denoted by

$$q_\mu = \sum_{v=0}^{\mu} \log \left(1 - \frac{1}{(1+\sigma+v\beta)^2} \right).$$

To investigate the behavior of the function C defined in (A-1), first consider $\sigma > 0$ fixed. Let $\beta = m^{-1}$, m being a positive integer. Then

$$-\infty < \log \left(1 - \frac{1}{(1+\sigma+v\beta)^2} \right) \leq \log \left(1 - \frac{1}{(1+\sigma+1)^2} \right) < 0 \quad (v = 0, 1, \dots, m),$$

and, hence, for $\mu > m$,

$$\begin{aligned} q_\mu &\leq m \log \left(1 - \frac{1}{(2+\sigma)^2} \right) + \sum_{v=m+1}^{\mu} \log \left(1 - \frac{1}{(1+v+\beta m^{-1})^2} \right) \\ &< \log \left(1 - \frac{1}{(2+\sigma)^2} \right) < 0. \end{aligned}$$

Therefore,

$$0 < \lim_{u \rightarrow \infty} e^{q_u} = \lim_{u \rightarrow \infty} \prod_{v=0}^u t_v = C\left(\frac{1}{m}, \sigma\right) < \exp\left(m \log\left(1 - \frac{1}{(2+\sigma)^2}\right)\right) \\ = \left(1 - \frac{1}{(2+\sigma)^2}\right)^m$$

The right-hand side can be made arbitrarily small if m is sufficiently large. In other words, for every fixed $\sigma > 0$, the infinite product $C(\beta, \sigma)$ will be arbitrarily small if $\beta > 0$ is sufficiently small, i.e., for every fixed $\sigma > 0$, it diverges to zero as $\beta \rightarrow 0$.

To investigate the behavior of $C(\beta, \sigma)$ for $\sigma > 0$ fixed and β large, use the partial products of (A-10),

$$p_u = \prod_{v=0}^u t_v = 1 - \frac{P}{\prod_{v=0}^u (1+\sigma+v\beta)^2}.$$

Relative to β , P is a polynomial of degree $2u$ with leading coefficient $u!$.

The denominator $\prod_{v=0}^u (1+\sigma+v\beta)^2$ is a polynomial in β also of degree $2u$ with leading coefficient $(1+\sigma)^2 u!$. Therefore, for every fixed $\sigma > 0$ and for every u ,

$$p_u = \prod_{v=0}^u t_v \rightarrow 1 - \frac{1}{(1+\sigma)^2} \text{ as } \beta \rightarrow \infty.$$

This means that, for every fixed $\sigma > 0$, the infinite product converges to $1 - (1+\sigma)^{-2}$, $0 < 1 - (1+\sigma)^{-2} < 1$, as $\beta \rightarrow \infty$.

From these results two preliminary conclusions are drawn:

(1) Since the function $C(\beta, \sigma) > 0$ in (A-1) can be made arbitrarily small for every fixed $\sigma > 0$ if $\beta > 0$ is sufficiently small, the function $h(\beta, \sigma)$ with $0 < A < 1$ will be negative for every fixed $\sigma > 0$ if $\beta > 0$ is sufficiently small and

(2) If $1 - (1+\sigma)^{-2} < A$, i.e., if

$$0 < \sigma < \sigma_A = \frac{1}{\sqrt{1-A}} - 1, \quad (\text{A-3})$$

$h(\beta, \sigma) < 0$ for every $\beta > 0$. On the other hand, $h(\beta, \sigma) > 0$ for every $\sigma > \sigma_A$ if β is sufficiently large.

Consequently, since h is a continuous function of β , observe that, for every fixed $\sigma > \sigma_A$, $h(\beta, \sigma) = 0$ has at least one positive root. Earlier it was observed that, for every fixed $\sigma > 0$, $h(\beta, \sigma) = 0$ has at most one root. It follows that, for every fixed $\sigma > \sigma_A$, $h(\beta, \sigma) = 0$ has exactly one positive root β .

With the established existence of at least one point $P^* = (\beta^*, \sigma^*)$, $\beta^* > 0$, $\sigma^* > \sigma_A$, such that $h(\beta^*, \sigma^*) = 0$, discussion of the equation $h(\beta, \sigma) = 0$ can be completed by means of the implicit function theorem. Its conditions are satisfied in some neighborhood of P^* : $h(\beta^*, \sigma^*) = 0$, $(h_\sigma)_{P^*} > 0$, $(h_\beta)_{P^*} > 0$, $h(\beta, \sigma)$, $h_\sigma(\beta, \sigma)$, $h_\beta(\beta, \sigma)$ being continuously differentiable in the domain $0 < \beta < +\infty$, $\sigma_A < \sigma < +\infty$. Consequently, there exists a closed interval $[\beta_1, \beta_2]$ such that $0 < \beta_1 < \beta^* < \beta_2$ and a one-valued continuous function $\sigma = \bar{\sigma}(\beta)$ such that $h(\beta, \bar{\sigma}(\beta)) = 0$ for every $\beta \in [\beta_1, \beta_2]$. The implicitly defined function $\bar{\sigma}(\beta)$ is even continuously differentiable in (β_1, β_2) . Its derivative is given by

$$\frac{d\bar{\sigma}(\beta)}{d\beta} = - \frac{h_\beta(\beta, \bar{\sigma}(\beta))}{h_\sigma(\beta, \bar{\sigma}(\beta))} < 0, \quad \beta \in (\beta_1, \beta_2),$$

i.e., $\bar{\sigma}(\beta)$ is a monotonically decreasing function of β for $\beta_1 \leq \beta \leq \beta_2$.

The domain of existence of the implicit function $\bar{\sigma}(\beta)$ can now be extended to all of $0 < \beta < +\infty$ by the following arguments.

Suppose $\bar{\sigma}(\beta)$ could not be continued to the right of some point $\bar{\beta} > 0$. Then there would exist a point \bar{P} on the line $\bar{\sigma} = \beta$ with coordinates $\bar{\beta}$ and $\bar{\sigma}$, $\sigma_A < \bar{\sigma} < +\infty$, such that every neighborhood of \bar{P} would contain infinitely many points P of the graph of the function $\bar{\sigma}(\beta)$ with abscissas $\beta < \bar{\beta}$. At each of these points $h(\beta, \sigma) = 0$. Then \bar{P} would be a limit point of such points P . Because of continuity of $h(\beta, \sigma)$ this would imply $h(\bar{\beta}, \bar{\sigma}) = 0$. But then the function $\bar{\sigma}(\beta)$ could be extended to the right of $\bar{\beta}$ by the original arguments. Analogous considerations apply for continuation to the left.

Since $h(\beta, \sigma) = 0$ has exactly one root $\beta > 0$ for every $\sigma > \sigma_A$, the range of $\bar{\sigma}(\beta)$ is the interval $(\sigma_A, +\infty)$ and, as a consequence of monotonicity, $\bar{\sigma}(\beta) \rightarrow +\infty$ as $\beta \rightarrow 0$, $\bar{\sigma}(\beta) \rightarrow \sigma_A$ as $\beta \rightarrow +\infty$.

APPENDIX B

Approximation of $v(u)$ by $v^*(u)$

Return to the function $v(u)$ defined in (8), replacing bM_1^{-1} by $\Gamma((1-p)8^{-1})/\Gamma((2-p)8^{-1})$ leaving, however, M_1b^{-1} in the argument of the ϕ -function unchanged for notational convenience. Then

$$v(u) = (1-p)u - \log \frac{1-p}{\beta} - (2-p) \log \Gamma\left(\frac{1-p}{\beta}\right) + (1-p) \log \Gamma\left(\frac{2-p}{\beta}\right) \\ + \log \phi\left[\frac{1-p}{\beta}, 1 + \frac{1-p}{\beta}; -\left(\frac{M_1}{b} e^u\right)^\beta\right]. \quad (B-1)$$

To approximate $v(u)$ the function $v^*(u)$ given in (11) was used with

$$v(0) = -\log \frac{1-p}{\beta} - (2-p) \log \Gamma\left(\frac{1-p}{\beta}\right) + (1-p) \log \Gamma\left(\frac{2-p}{\beta}\right) \\ + \log \phi\left[\frac{1-p}{\beta}, 1 + \frac{1-p}{\beta}; -\left(\frac{M_1}{b}\right)^\beta\right]$$

numerically to be determined by four-point Lagrange-Aitken interpolation from given points of the log cdf plot.

Subtraction of $v^*(u)$ from (B-1) results in

$$v(u) - v^*(u) = \rho + \log \phi(0) - \rho e^{\beta u}, \quad (B-2)$$

where, for notational convenience, $\phi(u)$ stands for the function ϕ as it appears in (B-1), $\phi(0)$ for its value at $u = 0$. Since $|v(u) - v^*(u)|$ must be small over a suitable u -interval and since $|v(u) - v^*(u)|$ must go to 0 as $u \rightarrow -\infty$, from (B-2) it can be seen that the two constants must satisfy the equation

$$\rho = \log \phi(0). \quad (B-3)$$

Now set $(1-p)8^{-1} = \alpha$, $(M_1b^{-1})^\beta = c$, and expand $\phi(u)$ into its power series [12;9.210.1],

$$\phi(u) = \phi(\alpha, 1+\alpha; -ce^{\beta u}) \\ = 1 - \frac{\alpha}{1+\alpha} \frac{1}{1!} ce^{\beta u} + \frac{\alpha}{2+\alpha} \frac{1}{2!} c^2 e^{2\beta u} - \frac{\alpha}{3+\alpha} \frac{1}{3!} c^3 e^{3\beta u} + \dots$$

If each of the exponential functions is expanded into its power series, the series for ϕ can be rearranged as follows:

$$\begin{aligned}\phi(u) = & 1 - \frac{\alpha}{1+\alpha} \frac{1}{1!} c + \frac{\alpha}{2+\alpha} \frac{1}{2!} c^2 - \frac{\alpha}{3+\alpha} \frac{1}{3!} c^3 + - \dots \\ & - \frac{\alpha}{1+\alpha} \frac{1}{1!} c \left[\frac{1}{1!} (\beta u) + \frac{1}{2!} (\beta u)^2 + \frac{1}{3!} (\beta u)^3 + \dots \right] \\ & + \frac{\alpha}{2+\alpha} \frac{1}{2!} c^2 \left[\frac{1}{1!} (2\beta u) + \frac{1}{2!} (2\beta u)^2 + \frac{1}{3!} (2\beta u)^3 + \dots \right] \\ & - \frac{\alpha}{3+\alpha} \frac{1}{3!} c^3 \left[\frac{1}{1!} (3\beta u) + \frac{1}{2!} (3\beta u)^2 + \frac{1}{3!} (3\beta u)^3 + \dots \right] + - \dots\end{aligned}$$

The series in the first row is equal to $\phi(0)$. The series in brackets in the v th following row is equal to $e^{v\beta u} - 1$. Therefore,

$$\phi(u) = \phi(0) - \sum_{v=1}^{\infty} (-1)^{v-1} \frac{\alpha}{v+\alpha} \frac{1}{v!} c^v (e^{v\beta u} - 1).$$

Denote the infinite series by $A(u)$. Then

$$\phi(u) = \phi(0)[1 - A(u)\phi^{-1}(0)] \quad (\text{B-4})$$

and consequently,

$$\log \phi(u) = \log \phi(0) + \log[1 - A(u)\phi^{-1}(0)].$$

Return with this expression to (B-2) and obtain

$$|v(u) - v^*(u)| = \left| \rho + \log[1 - A(u)\phi^{-1}(0)] - \rho e^{\beta u} \right|. \quad (\text{B-5})$$

The identity (B-4) shows that $1 - A(0)\phi^{-1}(0) = 1$. Furthermore, $0 < \phi(0) \leq \phi(u) < 1$ for $u \in (-\infty, 0)$, and $\phi(u) \uparrow 1$ as $u \downarrow -\infty$. Therefore, $1 - A(u)\phi^{-1}(0) \uparrow \phi^{-1}(0)$ as $u \downarrow -\infty$. Consequently,

$$1 \leq 1 - A(u)\phi^{-1}(0) < \phi^{-1}(0), \quad u \in (-\infty, 0],$$

which implies

$$0 \leq \log[1 - A(u)\phi^{-1}(0)] < -\log \phi(0), \quad u \in (-\infty, 0].$$

From (B-5) the following estimate now results:

$$|v(u) - v^*(u)| < |\rho - \log \phi(0) - \rho e^{\beta u}|, u \leq 0.$$

If (B-3) holds, this reduces to

$$|v(u) - v^*(u)| < |\rho| e^{\beta u}, u \leq 0,$$

and $v(0) - v^*(0) = 0$. Since $|v(u) - v^*(u)| \rightarrow 0$ as $u \rightarrow -\infty$, the maximum error in the approximation occurs at some $u_0 < 0$. Therefore

$$|v(u) - v^*(u)| \leq |\rho| e^{\beta u_0} \text{ uniformly in } u \leq 0.$$

The error over the interval $[0, u_k]$ is immaterial since the objective is to approximate $v(u)$ as $u \rightarrow -\infty$.

APPENDIX C

The Coefficient σ as a Function of R

Next the properties of the function $\sigma(R)$ defined in (14) together with (15) and (16) are investigated.

First of all, establish the fact that $\sigma(R) > 0$ for $R \in (0, +\infty)$. For two points $P_v = (u_v, v_v)$ ($v=1,2$) with $u_1 < u_2$ and the point $V_0 = (0, v(0))$

$$D = (u_1^2 + u_2^2)(a_1^2 + a_2^2) - (u_1 a_1 + u_2 a_2)^2 = (u_1 a_2 - u_2 a_1)^2$$

and, by induction,

$$D = A_{11} A_{22} - A_{12}^2 = \sum_{1 \leq \mu < \nu \leq k} (u_\mu a_\nu - u_\nu a_\mu)^2$$

for any number k of points $P_v = (u_v, v_v)$ with the abscissas ordered as above. Now look at the terms

$$u_\mu a_\nu - u_\nu a_\mu = u_\mu (e^{Ru_\nu} - 1) - u_\nu (e^{Ru_\mu} - 1), \quad 1 \leq \mu < \nu \leq k, \quad R > 0. \quad (C-1)$$

Set $u_\mu = x$, $u_\nu = y$, $x < y$, $x \neq 0$, $y \neq 0$. Then (C-1) changes into $x(e^y - 1) - y(e^x - 1)$ and $y = \alpha x$ leads to the function

$$f(x) = xg(x), \quad g(x) = e^{\alpha x} - 1 - \alpha(e^x - 1), \quad g(0) = 0, \quad (C-2)$$

$$g'(x) = \alpha e^x (e^{-(1-\alpha)x} - 1)$$

Distinguish the three possible cases.

1. $0 < x < y = \alpha x$, $1 < \alpha < +\infty$, so that $g'(x) > 0$, $x > 0$. Since $g(0) = 0$, $g(x) > 0$, $x > 0$ and, hence $f(x) > 0$, $x > 0$. Therefore $u_\mu a_\nu - u_\nu a_\mu > 0$, $0 < u_\mu < u_\nu$.

2. $x < 0 < y = \alpha x$, $-\infty < \alpha < 0$, $g'(x) < 0$. This and $g(0) = 0$ imply $g(x) > 0$, $f(x) < 0$, $x < 0$, and hence, $u_\mu a_\nu - u_\nu a_\mu < 0$, $u_\mu < 0 < u_\nu$.

3. $x < y = \alpha x < 0$, $0 < \alpha < 1$, $g'(x) > 0$. $g(0) = 0$ leads to $g(x) < 0$, $f(x) > 0$, $u_\mu a_\nu - u_\nu a_\mu > 0$, $0 < u_\mu < u_\nu < 0$.

Therefore, for (C-1)

$$u_\mu a_\nu - u_\nu a_\mu \begin{cases} < 0, & u_\mu u_\nu > 0, \\ < 0, & u_\mu u_\nu < 0. \end{cases} \quad (C-3)$$

Consequently, $D(\delta) > 0$, $\delta \in (0, +\infty)$.

Next, look at $D_1(\delta)$. For two points,

$$\begin{aligned} D_1 &= (u_1 c_1 + u_2 c_2)(a_1^2 + a_2^2) - (a_1 c_1 + a_2 c_2)(u_1 a_1 + u_2 a_2) \\ &= (u_1 a_2 - u_2 a_1)(c_1 a_2 - c_2 a_1) \end{aligned}$$

and, by induction,

$$D_1 = BA_{22} - CA_{12} = \sum_{1 \leq \mu < \nu \leq k} (u_\mu a_\nu - u_\nu a_\mu)(c_\mu a_\nu - c_\nu a_\mu).$$

Investigate the terms

$$\begin{aligned} c_\mu a_\nu - c_\nu a_\mu &= [v_\mu - v(0)](e^{R u_\nu - 1}) - [v_\nu - v(0)](e^{R u_\mu - 1}), \\ 1 \leq \mu < \nu \leq k, \quad R > 0. \end{aligned}$$

Division of (C-3) by $u_\mu u_\nu \neq 0$ results in

$$\frac{a_\nu}{u_\nu} - \frac{a_\mu}{u_\mu} > 0 \text{ in any case.} \quad (C-4)$$

Let $r_\mu = c_\mu u_\mu^{-1}$, $r_\nu = c_\nu u_\nu^{-1}$, and assume $0 < r_\nu < r_\mu$. Since r_κ may be interpreted as $\tan \theta_\kappa$, θ_κ being the angle between the horizontal positively oriented line through P_κ and the line through P_κ and $V_0 = (0, v(0))$, the last assumption implies concavity of the location of the three points P_μ , P_ν , and V_0 . Then, since $a_\kappa u_\kappa^{-1} > 0$, (C-4) implies

$$\frac{a_\nu}{u_\nu} - \frac{r_\nu}{r_\mu} \frac{a_\mu}{u_\mu} > 0 \text{ in any case.}$$

Multiplication by $u_\mu u_\nu r_\mu$ yields

$$u_\mu r_\mu a_\nu - u_\nu r_\nu a_\mu = c_\mu a_\nu - c_\nu a_\mu \quad \left\{ \begin{array}{l} > 0, \quad u_\mu u_\nu > 0, \\ < 0, \quad u_\mu u_\nu < 0. \end{array} \right. \quad (C-5)$$

On the basis of this inequality and (C-3), $D_1(\delta) > 0$, $\delta \in (0, +\infty)$, provided the points P_ν are concavely located. Consequently, under this concavity condition, the coefficient $\sigma(\delta)$ of the approximating function $v^*(u)$ is a positive function of $\delta > 0$.

The following remark is essential at this stage. In practical situations, all of the coordinates of the points P_ν ($\nu=1, \dots, k$) of a given empirical data set may not satisfy the inequality (C-5). Indeed, this is frequently the case.

However, violation of (C-5) will occur only for points P_v with v equal or close to 1, since the smoothing effect of the cumulative frequencies eliminates this occurrence for large values of v . In other words, if the points P_v are sufficiently smoothly located, D_1 and, consequently, σ , will still be positive. If, however, in a practical situation, D_1 should turn out to be always negative or zero, then this is a clear indication that the class (*) of distributions cannot be used for a data fit.

Turn now to the derivative of $\sigma(\beta)$ with respect to the parameter β . It is given by

$$\sigma' = D^{-2} \{ D[BA'_{22} - C'A_{12} - CA'_{12}] - D_1[A_{11}A'_{22} - 2A_{12}A'_{12}] \}, \quad (C-6)$$

$$A'_{12} = \sum_{v=1}^k u_v^2 b_v, \quad A'_{22} = 2 \sum_{v=1}^k u_v a_v b_v, \quad C' = \sum_{v=1}^k u_v b_v c_v, \quad b_v = e^{\beta u_v}.$$

Starting from $k = 2$ one can show by induction that

$$\begin{aligned} BA'_{22} - C'A_{12} - CA'_{12} &= \sum_{1 \leq \mu < v \leq k} [(u_\mu a_v - u_v a_\mu)(u_v b_v c_\mu - u_\mu b_\mu c_v) \\ &\quad + u_\mu u_v (c_\mu a_v - c_v a_\mu)(b_v - b_\mu)] \\ &= \sum_{1 \leq \mu < v \leq k} [(c_\mu u_v - c_v u_\mu)(u_\mu a_v b_\mu - u_v a_\mu b_v) \\ &\quad + 2u_\mu u_v (c_\mu a_v - c_v a_\mu)(b_v - b_\mu)] \end{aligned} \quad (C-7)$$

after addition and subtraction of $u_\mu u_v (c_\mu a_v - c_v a_\mu)(b_v - b_\mu)$, and

$$A_{11} A'_{22} - 2A_{12} A'_{12} = 2 \sum_{1 \leq \mu < v \leq k} u_\mu u_v (u_\mu a_v - u_v a_\mu)(b_v - b_\mu). \quad (C-8)$$

The derivative of $\sigma(\beta)$ can now be written as

$$\begin{aligned} \sigma' &= D^{-2} \{ [\sum (u_\mu a_v - u_v a_\mu)^2] [\sum (c_\mu u_v - c_v u_\mu)(u_\mu a_v b_\mu - u_v a_\mu b_v)] \\ &\quad + 2 [\sum (u_\mu a_v - u_v a_\mu)^2] [\sum u_\mu u_v (c_\mu a_v - c_v a_\mu)(b_v - b_\mu)] \\ &\quad - 2 [\sum (u_\mu a_v - u_v a_\mu)(c_\mu a_v - c_v a_\mu)] [\sum u_\mu u_v (u_\mu a_v - u_v a_\mu)(b_v - b_\mu)] \}, \end{aligned} \quad (C-9)$$

summations to be performed over $1 \leq \mu < v \leq k$. For simplicity, set

$$R_{\mu v} = u_\mu a_v - u_v a_\mu, \quad S_{\mu v} = c_\mu a_v - c_v a_\mu.$$

Then, using different subscript pairs for clearer distinction between the individual factors, write the second and third lines in this expression for σ' as follows:

$$2 \left[\sum R_{\mu\nu}^2 \right] \left[\sum u_{\kappa} u_{\lambda} S_{\kappa\lambda} (b_{\lambda} - b_{\kappa}) \right] \\ - 2 \left[\sum R_{\mu\nu} S_{\mu\nu} \right] \left[\sum u_{\kappa} u_{\lambda} R_{\kappa\lambda} (b_{\lambda} - b_{\kappa}) \right].$$

Those pairs of terms cancel for which the subscript pairs (μ, ν) and (κ, λ) are equal. The remaining terms are pairwise of the form

$$2R_{\mu\nu}^2 u_{\kappa} u_{\lambda} S_{\kappa\lambda} (b_{\lambda} - b_{\kappa}) - 2 R_{\mu\nu} S_{\mu\nu} u_{\kappa} u_{\lambda} R_{\kappa\lambda} (b_{\lambda} - b_{\kappa}), \quad (C-10)$$

$$(\mu, \nu) \neq (\kappa, \lambda).$$

The error equations (13) of the least squares approximation can be written as

$$\sigma u_{\nu} + \rho a_{\nu} - c_{\nu} = \varepsilon_{\nu} \quad (\nu=1, \dots, k).$$

Division by a_{ν} ($\neq 0$ since $u_{\nu} \neq 0$) results in

$$\sigma \frac{u_{\nu}}{a_{\nu}} + \rho - \frac{c_{\nu}}{a_{\nu}} = \frac{\varepsilon_{\nu}}{a_{\nu}}$$

and consequently,

$$\sigma \left(\frac{u_{\mu}}{a_{\mu}} - \frac{u_{\nu}}{a_{\nu}} \right) = \left(\frac{c_{\mu}}{a_{\mu}} - \frac{c_{\nu}}{a_{\nu}} \right) + \left(\frac{\varepsilon_{\mu}}{a_{\mu}} - \frac{\varepsilon_{\nu}}{a_{\nu}} \right).$$

Multiplication by $a_{\mu} a_{\nu}$ leads to

$$\sigma R_{\mu\nu} = S_{\mu\nu} + E_{\mu\nu} \quad (1 \leq \mu < \nu \leq k)$$

with $\varepsilon_{\mu} a_{\nu} - \varepsilon_{\nu} a_{\mu} = E_{\mu\nu}$. Then (C-10) changes into

$$2R_{\mu\nu}^2 u_{\kappa} u_{\lambda} (\sigma R_{\kappa\lambda} - E_{\kappa\lambda})(b_{\lambda} - b_{\kappa}) - 2R_{\mu\nu} (\sigma R_{\mu\nu} - E_{\mu\nu}) u_{\kappa} u_{\lambda} R_{\kappa\lambda} (b_{\lambda} - b_{\kappa}) \\ = 2R_{\mu\nu} u_{\kappa} u_{\lambda} (b_{\lambda} - b_{\kappa}) [E_{\mu\nu} R_{\kappa\lambda} - E_{\kappa\lambda} R_{\mu\nu}]. \quad (C-11)$$

If, for some $\delta > 0$, the points $P_{\nu} = (u_{\nu}, v_{\nu})$ should all be located on the graph of the function $v^*(u)$, then ε_{ν} would be zero for $\nu=1, \dots, k$. This would mean $E_{\mu\nu} = 0$ for $1 \leq \mu < \nu \leq k$, so that the terms in (C-11) would be zero. Hence, the derivative of $\sigma(\delta)$ would reduce from its general form (C-9) to

$$\sigma' = D^{-1} \sum_{1 \leq \mu < \nu \leq k} (c_{\mu} u_{\nu} - c_{\nu} u_{\mu}) (u_{\mu} a_{\nu} b_{\mu} - u_{\nu} a_{\mu} b_{\nu}).$$

In general, however, the approximation errors ϵ_v will not be zero. Then

$$\sigma' = D^{-1} \sum_{1 \leq \mu < \nu < k} (c_\mu u_\nu - c_\nu u_\mu)(u_\mu a_\nu b_\mu - u_\nu a_\mu b_\nu) + D^{-2} E \quad (C-12)$$

where E represents the sum of all terms (after cancellation) which are due to the ϵ_v 's not being zero.

Now establish inequalities for the factors in the sum in (C-12). First for $c_\mu u_\nu - c_\nu u_\mu$: if, as before, $r_\mu = c_\mu u_\mu^{-1}$, $r_\nu = c_\nu u_\nu^{-1}$, then $r_\mu - r_\nu > 0$ if the points P_μ , P_ν , V_0 satisfy the concavity condition. Multiplication of the last inequality by $u_\mu u_\nu$ leads to

$$c_\mu u_\nu - c_\nu u_\mu \begin{cases} > 0, & u_\mu u_\nu > 0, \\ < 0, & u_\mu u_\nu < 0. \end{cases} \quad (C-13)$$

Next look at

$$u_\mu a_\nu b_\mu - u_\nu a_\mu b_\nu = u_\mu (e^{\beta u_\nu - 1}) e^{\beta u_\mu - u_\nu} (e^{\beta u_\mu - 1}) e^{\beta u_\nu}.$$

Set $\beta u_\mu = x$, $\beta u_\nu = y$, $x < y$, $x \neq 0$, $y \neq 0$. Furthermore, set $y = \alpha x$, and obtain the function

$$f(x) = -x e^{-(1+\alpha)x} g(x), \quad g(x) = e^{-\alpha x} - 1 - \alpha(e^{-x} - 1).$$

With x replaced by $-x$, the function $g(x)$ appearing here is the same as that in (C-2). Therefore,

$$u_\mu a_\nu b_\mu - u_\nu a_\mu b_\nu \begin{cases} < 0, & u_\mu u_\nu > 0, \\ > 0, & u_\mu u_\nu < 0. \end{cases}$$

This inequality and (C-13) show that the sum in the expression (C-12) for σ' is certainly negative if the points $P_\nu = (u_\nu, v_\nu)$ are concavely located relative to each other with respect to the point $V_0 = (0, v(0))$ (or at least those for sufficiently large ν). Consequently, if the error term $D^{-2} |E|$ is sufficiently small, i.e., if the approximation of the points P_ν by the graph of the function $v^*(u)$ is sufficiently good, the derivative σ' of $\sigma(R)$ as given in (C-6) will still be negative.

An essential practical side result can be formulated on the basis of the last considerations. The class of distributions (*) may be used for an analytical fit of a given empirical statistical data set if the function $\sigma(R)$ is monotonically decreasing.

At the end of Appendix D this version of the applicability criterion shall be reformulated to obtain a practically more useful form.

This appendix shall be finished by an investigation of the limiting behavior of $\sigma(\beta)$ as $\beta \rightarrow +\infty$ and $\beta \rightarrow 0$. The function $\sigma(\beta)$ is defined by the expression

$$\sigma = \frac{BA_{22} - CA_{12}}{A_{11}A_{22} - A_{12}^2}, \quad (C-14)$$

the various terms being given in (16). As $\beta \rightarrow +\infty$, (since $u_1 < u_2 < \dots < u_{k-3} < 0 < u_{k-2} < u_{k-1} < u_k$),

$$BA_{22} - CA_{12} \sim (B - u_k c_k) a_k^2 + (\text{terms of lower order}),$$

$$A_{11}A_{22} - A_{12}^2 \sim (A_{11} - u_k^2) a_k^2 + (\text{terms of lower order}).$$

Therefore, provided σ' is negative,

$$\sigma + \sigma_\infty = \frac{B - u_k c_k}{A_{11} - u_k^2} = \frac{\sum_{v=1}^{k-1} u_v c_v}{\sum_{v=1}^{k-1} u_v^2} > 0, \quad \beta \rightarrow +\infty.$$

As $\beta \rightarrow 0$ the following is argued. Since the numerator and the denominator in the expression (C-14) for σ both go to zero as $\beta \rightarrow 0$, the Bernoulli-de L'Hospital rule applies. It is used three times. If, in the first step, the expressions (C-7) and (C-8) are used, the following expressions are derived in the third step for the numerator and denominator, respectively,

$$\begin{aligned} & \sum [u_\mu u_\nu (u_\nu b_\nu - u_\mu b_\mu) (u_\nu b_\nu c_\mu - u_\mu b_\mu c_\nu) + (\text{terms which go to } 0)] , \\ & 2 \sum [3u_\mu u_\nu (u_\nu b_\nu - u_\mu b_\mu) (b_\nu - b_\mu) + u_\mu u_\nu (u_\mu a_\nu - u_\nu a_\mu) (u_\nu^2 b_\nu - u_\mu^2 b_\mu)] . \end{aligned}$$

Clearly, each term in the second sum approaches zero as $\beta \rightarrow 0$. Since $b_\nu \rightarrow 0$ as $\beta \rightarrow 0$, the first term in the first sum approaches the positive value $u_\mu u_\nu (c_\mu u_\nu - c_\nu u_\mu) (u_\nu - u_\mu)$. Consequently, $\sigma \rightarrow +\infty$ as $\beta \rightarrow 0$.

Convergence of the Iteration Process

Two functional relations have been established:

$$\sigma = \bar{\sigma}(\beta), \quad \sigma = \sigma(\beta), \quad 0 < \beta < +\infty. \quad (D-1)$$

The first one is implicitly defined by the second moment equation $h(\beta, \sigma) = 0$; the second one has been derived from a least squares fit of given log cdf points. The properties of $\bar{\sigma}(\beta)$ and $\sigma(\beta)$ have been discussed in Appendixes A and C, respectively. The parameter determination problem has exactly one solution β_0 , $\sigma_0 = 1 - p_0$, b_0 , if and only if there exists exactly one value $\beta_0 > 0$ such that $\sigma(\beta_0) = \bar{\sigma}(\beta_0)$, i.e., geometrically speaking, if, and only if, the graphs of the two functions in (D-1) have exactly one point of intersection.

Suppose now that there is exactly one $\beta_0 > 0$ such that $\bar{\sigma}(\beta_0) = \sigma(\beta_0)$. Since $\bar{\sigma}(\beta)$ is not explicitly available, use instead of (D-1) the equivalent equations

$$h(\beta, \sigma) = 0, \quad \sigma = \sigma(\beta), \quad (D-2)$$

and solve them iteratively.

From the least squares fit by means of $v^*(u)$ with the starting value $\beta = 1$, exactly one number is obtained, $\sigma_1 = \sigma(1) = D_1(1)/D(1)$. Then solve the equation $h(\beta, \sigma_1) = 0$, its unique solution being β_1 and face the trichotomy $\beta_1 = 1$, $\beta_1 > 1$, $\beta_1 < 1$.

(a) If $\beta_1 = 1$, the iteration process through (D-2) produces the sequences $\{\sigma_v\}$ and $\{\beta_v\}$ with $\sigma_v = \sigma_1$, $\beta_v = 1$ ($v=1, 2, \dots$). Consequently, the solution of the system (D-2) is $\beta_0 = 1$, $\sigma_0 = 1 - p_0 = \sigma_1$.

(b) If $\beta_1 > 1$, sequences $\{\sigma_v\}$ and $\{\beta_v\}$ are obtained for which not all elements are equal. If the function $\sigma(\beta)$ is monotonically decreasing (Appendix C), i.e., if the given data are not "ill-conditioned," in the second iteration step a number $\sigma_2 = \sigma(\beta_1) < \sigma_1 = \sigma(1)$, is obtained. Since $h(\beta_1, \sigma_1) = 0$, $h(\beta_1, \sigma_2) < 0$ (Appendix C). Therefore, the root β_2 of $h(\beta, \sigma_2) = 0$ satisfies the inequality $\beta_2 > \beta_1$. These arguments apply in each of the subsequent iteration steps. Since it was assumed that there exists only one value β_0 for which $\sigma(\beta_0) = \bar{\sigma}(\beta_0)$, the sequence $\{\beta_v\}$ converges to β_0 , $\beta_v + \beta_0 > 1$ as $v \rightarrow +\infty$, and the sequence $\{\sigma_v\}$ converges, $\sigma_v + \sigma_0 = 1 - p_0$ as $v \rightarrow +\infty$.

(c) If $\beta_1 < 1$ analogous arguments apply to establish convergence of the sequences $\{\beta_v\}$ and $\{\sigma_v\}$ to unique limits $\beta_0 < 1$, $\sigma_0 = 1 - p_0$, respectively.

In practice, of course, the iteration process is stopped whenever a desired accuracy has been reached, and the last β -value is taken as β_0 .

Convergence of the iteration process represents the ultimate practical test for the applicability of the class of distributions (*). If the iteration process does not converge, this is, in fact, an indication that the given empirical data set is ill-conditioned and that the class (*) should not be applied.

DISTRIBUTION

	<u>Copies</u>
Director	
US Army Research Office	
ATTN: SLCRO-PH	1
SLCRO-ZC	1
P. O. Box 12211	
Research Triangle Park, NC 27709-2211	
Headquarters	
Department of the Army	
ATTN: DAMA-ARR	1
Washington, DC 20310-0632	
OUSDRE	
ATTN: Dr. Ted Berlincourt	1
The Pentagon	
Washington, DC 20310-0632	
US Army Materiel System Analysis Activity	
ATTN: AMXSY-MP	1
Aberdeen Proving Ground, MD 21005	
IIT Research Institute	
ATTN: GACIAC	1
10 W. 35th Street	
Chicago, IL 60616	
AMSMI-RD, Dr. McCorkle	1
Dr. Rhoades	1
-RD-RE, Dr. R. Hartman	1
Dr. J. Bennett	1
Dr. S. L. Lehnigk	10
Mr. H. P. Dudel	10
-RD-CS-T, Record	1
-RD-CS-R, Reference	15
-GC-IP, Mr. Bush	1